A method for quantifying generalized synchrony in brain activity

B. Richard McAteer
Carleton University

Copyright © 2003, B. Richard McAteer

National Library
of Canada

Bibliothèque nationale
du Canada

Acquisitions and
Bibliographic Services

Acquisitions et
services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 0-612-93986-3

Our file Notre référence

ISBN: 0-612-93986-3

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this dissertation.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de ce manuscrit.

While these forms may be included in the document page count, their removal does not represent any loss of content from the dissertation.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

PAGINATION ERROR.

ERREUR DE PAGINATION.

TEXT COMPLETE.

LE TEXTE EST COMPLET.

Abstract

The development of new methods of quantifying brain dynamics may shed light on the mechanisms of epilepsy and increase the ability to predict seizure. Many attempts to predict seizure were based, with limited success, on measuring complexity of a signal (e.g. correlation dimension). Here a non-linear method has been developed, using a nearest-neighbour approach to quantify the level of predictability ($Dv_{\max}$) between signals from different brain regions. The measure is non-symmetric, allowing for the detection of unidirectional influences, and is insensitive to the phase of periodic signals.

A program was written, using LabVIEW, to evaluate $Dv_{\max}$ between four simultaneously recorded signals. The program was tested on model systems with known properties, and on recordings from hippocampal and amygdaloid electrodes in rats bred to be prone (Fast) or resistant (Slow) to kindling – an animal model of epilepsy. The minimum embedding dimension of brain activity was lower following kindling. Furthermore, Fast and Slow rats had significantly different $Dv_{\max}$ before kindling, and these differences disappeared after kindling.

Acknowledgements

I would like to thank Tina Price-O'Dea for her work in the surgeries, depth electrode recordings, behavioural scoring and kindling of the rats; Pavel Gromiek for having classified sections of those recordings by exhibited behaviour and compiled them in separate files, the McIntyre Lab for allowing me to use their computer and lab space, Bruce Hutcheon for his guidance in this project and the numerous philosophical, trivial, and encyclopaedic discourses we have had on otherwise unproductive days, and for being my carrot-and-stick.

I would also like to thank my wife Sara for her love and support through my various scholastic endeavours, and my daughter Morgan, who makes all worthwhile.

Table of Contents

Title page	i
Acceptance form	ii
Abstract	iii
Acknowledgements	iv
Table of contents	v
List of tables	vi
List of Figures	vii
List of Abbreviations and Symbols	viii
Introduction	1
Methods	15
Results	30
Discussion	65
References	94
Appendix A: Data	103
Appendix B: Functions	171

List of Tables

Table 1	56
---------	----

List of Figures

Figure 1	7
Figure 2	20
Figure 3	20
Figure 4	25
Figure 5	25
Figure 6	39
Figure 7	42
Figure 8	42
Figure 9	45
Figure 10	45
Figure 11	49
Figure 12	49
Figure 13	51
Figure 14	51
Figure 15	55
Figure 16	55
Figure 17	59
Figure 18	59
Figure 19	62
Figure 20	63
Figure 21	64

List of Abbreviations and Symbols

A_{Tol}	Maximum ratio of neighbour distance to attractor size for true neighbours
Δt	Time shift applied to reference attractor for the divination function
d_A	Measure of attractor size
$d_m(x,y)$	Distance from x to y in m dimensions
D_A	Dimension of attractor
$Dv(X,Y)$	Divination coefficient for X predicted from Y
$Dv(X,Y; \Delta t)$	Divination function for X predicted from Y with time shift Δt
$Dv_{max}(X,Y)$	Maximum of the divination function for X predicted from Y
EEG	Electroencephalogram
FNNP	False nearest neighbour percentage
k	Number of neighbours
m	Embedding dimension
n	Time index, also t_n
N	Number of points, number of embedded points
$n(\mathbf{x}(t),i)$	i^{th} nearest neighbour of $\mathbf{x}(t)$
R_{Tol}	Maximum ratio of nearest neighbour distance from dimension $m+1$ to dimension m for true neighbours
rnd	Referred neighbour distance
$r(\mathbf{x}(t),i,Y)$	i^{th} neighbour referred from Y at time t
s	Sampling, the fraction of data points that are embedded
SNR	Signal to noise ratio
t	Time, t_n when discrete, being the time for the n^{th} point

List of abbreviations and Symbols cont'd

τ	Delay for the embedding
τ_{Theiler}	Minimum difference in time between a point and point eligible for consideration as neighbours.
tnd	True neighbour distance
$x(t)$	A signal x , evaluated at time t . Also the whole signal.
$\mathbf{x}_m(t)$	The embedded point (vector) formed from the signal x at time t .
X	The set of all embedded points (vectors) of $x(t)$

Introduction

In this thesis I have created programs to implement a technique for detecting nonlinear correlations between brain regions, for use with electroencephalograms (EEGs) and other records of electrical activity. The basis of the method is to identify attractors for the dynamics underlying the electrical signal recorded from different brain regions and then determine how much a knowledge of the properties of the attractor in one area helps in reconstructing the properties of the attractors in other areas, similar to the *generalized synchronization* proposed by Rulkov et al. (1995). A similar method, *nonlinear mutual predictability* (Le Van Quyen et al., 1998) was used to examine the recordings of pairs of ipsilateral hippocampal depth electrodes during human temporal lobe seizure; the method I have developed is based on the method outlined in Arnhold et al. (1999), and is additionally able to detect time lags between structures.

Despite the incredible complexity of the brain, with its 10^{10} neurons, and 10^{15} connecting synapses (King, 1991), people have tried for ages to come to terms with its activities. Its purpose has been misunderstood over the millennia, having been thought of as waste by the Egyptians - who discarded it yet kept the liver, stomach, lungs and intestines - to Descartes, who thought that the pineal gland was the seat of the soul and that the muscles moved due to the application of pressurized fluid from the ventricles. Only in the late 19th century did a technique become available to study the function of the brain, and it was not until 1924 that Hans Berger first recorded a human EEG (Malmivuo, 1999). EEGs are linked to mental states – the pattern of an EEG varies with activity (King, 1991), and one of the most noticeable changes is the frequency, with slow wave sleep showing activity at a low frequency (2-3Hz), while alpha waves in the awake brain

are around 10Hz (Nunez, 1995). Large amplitudes on the EEG represent the summed fields of many neurons; thus synchrony of neurons is one of the most obvious features on the EEG and the manner in which the brain synchronizes such a large network is of great importance.

Like many complex systems, we can learn the most about the brain's dynamics when it malfunctions. One type of malfunction that has been studied in depth is epilepsy, which affects approximately 1% of the population, and has distinctive EEG correlates (Nunez, 1995). The large-amplitude paroxysmal activity characteristic of the EEG during seizure reflects large numbers of neurons firing in tight synchrony. This striking feature of the EEG no doubt fueled attempts to understand the etiology of seizures through methods that examine the EEG for changes in synchrony. However, large-scale synchrony is not exclusively associated with seizure. Normal brain function includes many types of rhythmic activity, from the oscillations of slow wave sleep to the fast oscillations of alert activity, and synchronies develop in these oscillations based on task performance, motivation and sensory stimuli (Steriade, 2001). The paroxysmal activity in epilepsy could very well be rooted in a malfunction of networks of over-synchronous neurons.

The ability of some epilepsy patients to anticipate their seizures via prodromes (Rajna et al., 1997) is an indication of changes in function that presage seizure, and had noted changes in cerebral blood flow prior to seizure, indicating that physiological processes were at work (Le Van Quyen et al., 2001). The changes occurring before the clinical manifestation have been the target of many investigations, and the detection of this pre-seizure period has potential benefits in terms of possible interventions, as well as giving us the ability to study the generation of seizure and understand its development.

Seizure prediction has thus been the goal of many researchers, with varying degrees of success; due to the prevalence of this type of research in recent epilepsy work I will be using examples of methods from seizure prediction in the following sections to compare to the method I have developed. The types of analyses are easily divided into those relying on linear methods and those relying on nonlinear methods, and I will now examine representative methods from both camps with respect to their ability to detect the changes preceding seizure, and thus of brain dynamics.

Linear methods

Much of contemporary mathematical analysis of the EEG is derived from the techniques used in physics and engineering, and the simplest approach to analysis of signals (such as EEG) is to treat the signal as coming from a linear system. Linear systems follow the principle that small changes in input result in small changes in output; a linear method interprets all of the regularities in a system as being due to linear correlations (Kantz and Schreiber, 1999). Linear systems can only result in irregular outputs when driven by irregular inputs - in the absence of perturbations, linear systems either oscillate periodically or grow or shrink exponentially (Kantz and Schreiber, 1999). If the brain is simply a linear system analyses based on linear methods should be able to predict the onset of seizure accurately.

One of the most important properties of linear systems is that the summed output resulting from two signals is the same as the output if the two signals were summed initially. This allows the use of Fourier transforms to decompose the signal into its component frequencies and measure the amplitude and phase of these Fourier

frequencies. The power spectrum - the squared modulus of the Fourier transform (Kantz and Shreiber, 1999) - is a tool used in many linear analyses and is a standard way to examine the dominant frequencies in a signal. The changes in the strength of frequencies that are seizure related can be used to determine the onset of seizure, but with limited effectiveness.

The use of power spectra (Siegel et al., 1982) allowed for some degree of prediction of the onset of epileptic seizure, but the prediction times for these methods were only in the realm of a few seconds (Lehnertz and Elger, 1998; Le Van Quyen and Martinerie, 1999). A linear auto-regressive model used by Rogowski et al. (1981) had a prediction time of several seconds, but given that the cerebral blood flow changes as early as 10 minutes before seizure (Weinand et al., 1997) and that premonitory symptoms are reported by many subjects often preceding the seizure by five minutes or more (Rajna et al., 1997) these methods fall short of detecting the origin of the disruption.

Nonlinear methods

Nonlinear methods are simply those methods that are not linear, thus they lack the additive properties of the linear methods, and are not subject to the constraint that a small change in the input results in a small change in the output. The brain has many features which suggest that it may be nonlinear in function, action potentials are a nonlinear phenomenon; two inputs to a neuron can fail to sum to an action potential, yet slightly larger inputs or simply more closely timed inputs may result in the release of an action potential. This type of large effect from a small change in the input is a nonlinear behaviour, and a system composed of nonlinear units can, with very few exceptions, be

presumed to be nonlinear; thus we would expect to see better performance in seizure prediction with nonlinear methods.

Additional support for the theory that the brain is nonlinear arises from the fact that unlike a linear model, nonlinear systems can exhibit complex quasi-periodic activity on simple inputs, and can have stable limit cycle oscillations (Glass, 2001). While linear systems can have periodic behaviour, they are regular periodic oscillators, not the quasi-periodic oscillators that can evolve from nonlinear systems. A good example of a periodic oscillator is a pendulum – the pendulum will swing on the exact same arc so long as it is undisturbed, but when an outside force is applied its behaviour changes, and it will never return to its original arc unless forced there by another outside force. A nonlinear oscillator with a limit cycle has a set of states from which it can return to the original oscillation, called the *basin of attraction* (Kantz and Schreiber, 1999); within that range of states the dynamics will pull the system back to the same limit cycle. Complex systems can have more than one limit cycle, and through perturbation can be forced into the basin of attraction of a different limit cycle; this is a good model for the manner in which the brain can enter and exit seizure states, and suggests that nonlinear systems may be a better model for the brain's dynamics.

Rather than try to examine all of the details of systems with this type of periodicity we can examine a simpler entity called the *attractor* – the set of states to which the system tends. The attractor is essentially the limit cycle to which the system will return when undisturbed; we can think of the brain as having attractors that tend to keep activity in certain patterns. The epileptic seizure thus has its own attractor, and when we enter the basin of attraction for a seizure we fall toward that attractor. In order to study the

attractor of a system we embed it in *state space*, a process that assigns the state at any time to a point in a multi-dimensional space to evince the ways in which the state can change (Kantz and Schreiber, 1999).

A common way to embed a time series is through the *method of delays* (Takens, 1981). The method of delays reconstructs the original dynamics of the EEG based on the time series; given the huge number of neurons and the interconnections between them it is not feasible to measure what is occurring at every neuron – rather, a celebrated theorem by Takens (1981) explains that if we measure a single variable accurately over time we can describe the underlying dynamics. The simplest way to do this is to plot the value at any time against the value a set time (τ) earlier. While noise occupies this 2-dimensional space at random, organized systems occupy a subset of the state space (Figure 1), and one of the nonlinear measures associated with this is the dimension of the state space necessary for the reconstruction of the attractor.

With the attractor represented in state space the points (vectors) represent the system dynamics leading up to that time index – they are a snapshot of the recent changes. Points that are near one another represent similar states of the system with neighbouring points being the most similar. The attractor's orbits are states through which the system passes regularly, and similarity of dynamics can be inspected geometrically on the attractor.

Figure 1. Comparison between the embedding of uniform white noise (left) and a recording from a depth electrode in rat hippocampus during seizure (right). The original signal is underneath each embedding; it is clear that there is an order to the seizure that is not present in the noise, and the path through state space is an estimate of the attractor. The paths in the embedded seizure signal only appear to cross due to the presentation as a 2 dimensional projection; in the appropriate embedding dimension they would not meet.

Much of the work that has been done to analyse the EEG is in terms of the correlation dimension, the minimum number of variables that are needed to represent the attractor of the underlying dynamics (Le Van Quyen et al., 2001). If every neuron were free to behave independently the dimension of the reconstructed state space of the EEG would be at least the number of neurons involved, but the correlation dimension of the brain has been estimated to be about 10, with lower dimensions for slow wave sleep (4-5) and epilepsy (3-4) (King, 1991). The very low dimensions involved suggest that the relationships between neurons constrain what type of activity occurs; the lower dimension that is found with slow wave sleep and epilepsy is then representative of a less complex signal. A drop in signal complexity is thus a symptom of the change in dynamics, which can be observed in the delay-embedded points.

The finding that the dimension of the attractor changes when different activities are performed suggests a change in nonlinear dynamics occurs. The drop in minimum embedding dimension at the start of a seizure, most noticeably in the seizure focus (Lehnertz and Elger, 1995), shows that this is sensitive to the changes that occur in seizure. Given the existence of a pre-seizure period it is natural to try to apply the methods based on nonlinear dynamical change to seizure anticipation. Several methods of nonlinear analysis have attempted to quantify this change in dynamics, including neuronal complexity loss (Lehnertz and Elger, 1995, 1998), similarity indices (Le Van Quyen et al., 1999, 2001, 2001b; Martinerie et al., 1998; Navarro et al., 2002), Lyapunov exponents (Moser et al., 1999) and mutual prediction (Le Van Quyen et al., 1998). While the anticipation times tend to be in the range of a few seconds using the linear methods

mentioned above, non-linear methods have anticipation times in the 2-11 minute range (Elger and Lehnertz, 1998; Lehnertz and Elger, 1998; Le Van Quyen et al., 1999, 2000, 2001; Martinerie et al. 1998; Navarro et al., 2002).

The existence of a pre-seizure state has been supported by numerous studies (Martinerie et al., 1998; Mormann et al., 2003; Navarro et al., 2002) and the nonlinear methods appear to be more sensitive to these changes, finding that not only are the anticipation times longer with the nonlinear methods, but that nonlinear methods can find significance in cases where linear analyses fail; for example, traditional spectral analysis of the EEG was less than 20% as sensitive to pre-seizure changes when compared to nonlinear methods (Navarro et al., 2002). The effectiveness of the nonlinear methods has led some researchers to feel that there is little doubt that the dynamics of the brain are nonlinear (Casdagli et al., 1997).

Multiple signal analysis

Since the spread of an epileptic seizure is in some cases thought to originate from a focus and spread to neighbouring regions it is sensible to try to find changes in synchronicity within the brain. If we imagine a seizure spreading from the epileptic focus by entraining other brain regions we can see why examining the building synchronies between previously independent brain regions would be a more sensitive technique. The techniques comparing simultaneously recorded signals from different regions of the brain (or from different cells) take advantage of a greater amount of information than do those techniques that measure single values – the overall correlation dimension could stay constant while the relationships between brain regions is

undergoing changes. While the embedding theory by Takens (1981) shows that we can reconstruct the dynamics of a system by observing a single variable, this technique assumes an infinite data set and no measurement error. Both linear and nonlinear methods can be applied to multiple signals – the goals of these studies vary from attempts to improve anticipation times in epilepsy (Mormann et al., 2003), to attempts to understand the synchronies in the brain at the cellular (Netoff and Schiff, 2002) and structural levels (Le Van Quyen et al., 1998; Meeren et al., 2002). I will review the improvements in anticipation times that result from these studies.

At the cellular level, it has been shown that cells in the hippocampus undergo a decrease in synchrony with other nearby (30-200 μm) cells, as measured in cross-correlation, mutual information and phase-correlation, while the cross-correlation of the cells against an extracellular electrode showed increased cross-correlation during the seizure (Netoff and Schiff, 2002). This finding is odd, given the apparent increase in synchrony during seizure, but may have to do with a loss of communication between organised circuits and the loss of ordinary patterns of activity. The notion that communication may be lost between organized groups of cells during the development of seizures is supported by pre-ictal drops in synchrony (Mormann et al. 2000, 2003, 2003b), and the levels of synchrony rise to normal after the seizure. The more recent paper by Mormann et al. (2003b) using a phase coherence approach has demonstrated anticipation times in the range of roughly one hour; this is a huge improvement on the results we have seen with either the non-linear methods or the linear methods. An important distinction between this method and other linear methods is that the phase coherence is not amplitude sensitive, but is based on computing a phase for the

frequencies that make up the signal; it is still a linear method, but the amount of useful information clearly increases, as the anticipation times increase drastically when concurrent recordings are compared.

By using the information about synchronies within the brain rather than looking at an overall similarity index or loss of neuronal complexity there is a greater ability to find synchronies between brain regions which would be obscured by univariate methods, and to understand how recruiting may take place during the initiation and spread of a seizure. The epileptic focus shows a more pronounced loss of complexity (Lehnertz et al., 1999), and appears to recruit other regions into the paroxysmal activity. In order to study the spread of seizure it is important to determine not just whether regions of the brain are synchronous, but whether one region drives another region to epileptiform activity. In a recent study of seizure development a grid of cortical electrodes was implanted in the cortex of rats suffering from absence epilepsy, and non-linear correlations were used to determine the strength of the relationships between electrode sites (Meeren et al., 2002). Importantly, the non-linear correlation was calculated for a range of lag values, and the maximum was chosen to determine the direction of any lead-lag associations. During seizure the association between the leading site and other sites became stronger, and the lead-lag effect intensified; both of these effects dropped off quickly at the end of seizure. The distance between electrodes was in general inversely related to the strength of the association and directly linked to the lag time, suggesting that the lagging regions are driven by the leading region.

The ability to detect the correlations between brain regions is important in characterizing the manner in which seizures develop, and the lag-based method of

determining whether one region leads another allows for the possibility of locating driving populations. All of the methods mentioned thus far are symmetric; a comparison of A vs. B gives the same result as B vs. A, but when considering correlations between attractors we do not need to limit ourselves to a symmetric relationship. Especially when one signal drives several others, or when the signals all have correlations to one another, the strength of prediction may be greater in the direction of the driving force than vice versa. A nonlinear measure termed *nonlinear mutual predictability* is one measure of the strength of relationship between two signals that does not have the property of symmetry (Le Van Quyen et al., 1998). Nonlinear mutual predictability relies on embedding the EEG in phase space according to the method of delays, and making a prediction of a point's position at time $t+1$ based on the value of the mutual neighbours to the point at time t , and allowing those neighbours to step forward in time by one unit. The averaged predicted change is compared with the actual change, and a statistic is generated quantifying the predictive power that one signal has for another. These comparisons can be greater in one direction than the other, which the Le Van Quyen group interpret as showing the direction of a driving force.

The method described by Arnhold et al. (1999) is also a nonsymmetric method, and in their paper they suggest a way to generalize the method to a time-shifted method, much like the generalization of the cross-correlation done by Meeren et al. (2002). The program I have developed uses this generalized form, and is sensitive to temporal lag between signals, direction of driving force and capable of detecting both linear and generalized synchronies.

This value, which for ease of reference I am dubbing the divination coefficient (as it is unnamed in Arnhold et al., and simply referred to as “a measure of interdependence”), will then reflect the degree to which knowing the state of one signal predicts the state of the other signal. The method is nonlinear, and relies on a delay embedding to compare the states of the signals by using the mutual neighbourhoods.

Having developed a program to identify small changes in the level of synchrony between regions of the brain, I have chosen to use the method to look for changes in the way in which regions of the brain synchronized before, during and after kindling. Since kindling is a model of epilepsy in the rat, we can expect that the nonlinear methods that work for detecting the changes in synchrony before seizure will be effective at detecting the changes in the underlying dynamics during kindling.

Kindling and nonlinear correlation

Kindling is a process by which an animal can have the threshold for stimulated seizure lowered by repeated stimulation (Racine et al., 1999). This phenomenon was first noticed by Goddard (1967) with daily stimulus of the amygdala resulting in the development of convulsive seizures. Repeated kindling can result in spontaneous convulsive seizures, and these can be developed with no obvious brain damage (Michael et al., 1998). The development of Fast and Slow strains of rat (Racine et al., 1999) which respond differently to kindling is an indication that some of the systems are under genetic control, and provide us with some interesting opportunities to understand how seizures may develop.

The Fast and Slow rats differ in their ability to undergo kindling, the severity and duration of the convulsions and of the after discharge (Racine et al., 1999). I propose to measure the general synchrony, as measured by the maximum of the divination function (Dv_{max}), between four regions of the brain before and after kindling, in the Fast and Slow strains of rat, to determine whether the changes in behaviour are associated with changes in the levels of synchrony, and whether the strains differ in this measure. The levels of synchrony between regions will also be measured for the kindling sequence, since even sub-threshold stimulation can cause changes in the observed EEG, and these smaller stimuli may show changes in synchrony, and reveal differences in the kindling sequence between the Fast and Slow strains.

Methods

Overview

For my thesis I have developed a program to measure a form of predictability between signals, the *divination coefficient*, which is computed in a manner similar to that set out by Arnhold et al. (1999), with a few differences to allow for the generalization to the *divination function*. The word divination is defined by the Merriam-Webster collegiate dictionary as “the art or practice that seeks to foresee or foretell future events or discover hidden knowledge usually by the interpretation of omens or by the aid of supernatural powers”, not a wholly inaccurate word for this type of prediction between embedded time series.

As will be explained more fully below, to calculate the divination coefficient, the signals to be compared, $x(t)$ and $y(t)$, are embedded using the method of delays (Takens, 1981) to produce the embedded sets X, Y . To predict X from Y we first locate a point $\mathbf{x}(t_0)$ in X . This will become our sample point. The average distance from the point $\mathbf{x}(t_0)$ to its k nearest neighbours is calculated (this is the true neighbour distance, *tnd*). A corresponding reference point $\mathbf{y}(t_0)$ is found in the reference set Y , its k closest neighbours are found, and the points in X with time indices matching these neighbours are identified. The points in X so identified are not nearest neighbours of $\mathbf{x}(t_0)$, but *referred neighbours*, as they represent the neighbourhood in X predicted from Y . The average distance from $\mathbf{x}(t_0)$ to these referred neighbours is calculated (referred neighbour distance, *rnd*), and the ratio of the true neighbour distance to the referred neighbour distance is calculated (*tnd/rnd*). The average of this value over all sample points $\mathbf{x}(t)$ is the divination coefficient, $Dv(X, Y)$. This is read as the divination value of X predicted from Y .

When there exists a generalized synchrony between two signals (Rulkov et al., 1995) $Dv(X,Y)$ value will be near 1. The value of $Dv(X,Y)$ is constrained between 0 and 1, as it is a ratio of positive values and the referred neighbours cannot be closer than the nearest neighbours. If a generalized synchrony exists it need not be bidirectional, but can consist of a driver and a response system, and using this system the predictability of X from Y ($Dv(X,Y)$) is not necessarily equal to the predictability of Y from X ($Dv(Y,X)$).

A system exhibiting generalized synchrony may have lags between signals that may be important functionally. We will evaluate the divination coefficient for a set of time lags, such that the coefficient is calculated for a point at time t , predicting the other attractor at times $t+\Delta t$, where Δt is the amount of simulated time lag between the signals. This produces divination values as a function of t , and analogous to the cross-correlation function, we can call this the *divination function* between the time series. By taking the maximum value of the divination function among the choices for Δt we find the best prediction lag, which is the estimated lag time.

Embedding of time series

The embedding of the time series is done using delay embedding (Takens, 1981) and creates an embedded time series as follows. For a time series $x(t)$ the delay reconstruction with dimension m and delay τ is given by

$$\mathbf{x}_m(t_n) = (x(t_n), x(t_{n-\tau}), \dots, x(t_{n-(m-1)\tau}))$$

This embedding procedure will encode all but $(m-1)\tau$ points, so when starting with N measurements, $N - (m-1)\tau$ coordinates of dimension m are produced. We can call the set of all points of the embedded time series X , i.e.

$$X = \{x_m(t_n) | 0 < n < N - (m-1)\tau, n \in \mathbb{N}\}$$

The embedded points can then be thought of as forming a path in m -dimensional space (which we will call state space) as the points have an order imposed on them in terms of the time index t_n . As we see from the description for the embedding above, there are 2 parameters to be specified, the dimension, m , and the delay, τ . The set X thus depends on the choice of these parameters. Although the notation $\mathbf{x}_m(t_n)$ is most precise as it specifies the time index n and the embedding dimension m , we can write $\mathbf{x}(t)$ to represent the embedded point in X if the dimension and index are not needed.

If a time series results from a system that is self organizing we expect the behaviour of such a system to have a low dimensional attractor; this dimension is in some ways a measure of the complexity of the system's behaviour – simple signals such as a sine wave can be fully expressed in 2 dimensions, while more complex signals require more dimensions to prevent ambiguity. According to a theorem due to Takens (1981), the single observed value $x(t)$ is a projection of state space onto a single dimension, and by choosing a dimension m greater than twice the attractor dimension we recreate the topology of the attractor. This limit $m > 2 D_A$ is not a minimum for all attractors, it is the lowest value of m that will work for all attractors of dimension D_A ; the Lorenz attractor has $D_A=2.06$, but can be embedded in $m=3$ without ambiguity (Kennel et al, 1992). Since the time involved in computation increases with the dimension used we would like to use the smallest dimension capable of expressing the system's dynamics, and as there is not a known value of the dimension of the brain's state, we need to evaluate the dimension based on the data. In addition, since no improvement in resolution of the dynamics occur once we have reached the minimum embedding dimension, the effect of noise is

increased when we use more dimensions than necessary, as the noise continues to increase the spread of the data as we increase the dimension.

Embedding dimension and neighbours

The method used to determine an appropriate embedding dimension is that of false neighbours (Kennel et al, 1992). A neighbour is simply the point that is closest to a given sample point, as we are interested in neighbours within this space a metric must be used to measure distances between points; a metric is a function used to measure the distance between points x and y . Some example metrics would be the Euclidean metric which we use in normal activities, the Manhattan metric which is used to describe distances on a square grid or on city blocks (and is alternately the block or taxicab metric) and the maximum metric which is a measure of the greatest difference between coordinates. The metric we chose to use is the standard Euclidean metric, the distance between two points x and y in m dimensions being

$$d_m(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_m - y_m)^2}$$

$$x = (x_1, x_2, \dots, x_m)$$

$$y = (y_1, y_2, \dots, y_m)$$

The Euclidean metric was chosen due to the familiarity of the metric and should be easily substituted with another metric within the program; as well it is a metric that most are familiar with, and lends itself to easy interpretation. For the purpose of finding neighbours the squared Euclidean distance is equivalent, as both provide identical orderings of neighbours, and the squared distance takes fewer steps to calculate. The i^{th} nearest neighbour $n(\mathbf{x}(t), i)$ is the point i^{th} closest to $\mathbf{x}(t)$, i.e. $n(\mathbf{x}(t_0), 1)$ is the point closest to $\mathbf{x}(t_0)$, while $n(\mathbf{x}(t_0), 2)$ is the next closest point, and for simplicity, $n(\mathbf{x}(t), 1)$ can be

shortened to $n(\mathbf{x}(t))$ for the case of the nearest neighbour. When the dimension in which we embed is below the minimum dimension for the attractor we have a projection of the dynamics into a lower dimensional space, much like the 2-dimensional shadow cast by a 3-dimensional object. In such projections points can appear to be near one another, while being distant in the higher dimension.

In Figure 2 we see the embedding of a sine wave in two dimensions, and the same sine wave embedded in one dimension. Points 1 and 2 are close in both the one and 2-dimensional embeddings, while point 3 seems close to both 1 and 2 in a 1-dimensional embedding, but is in fact quite far in a 2-dimensional embedding. This can be seen again in a more complex situation; Figure 3 shows three 2-dimensional projections of the 3-dimensional embedding of 1000 points of the Lorenz attractor; again we see that in some two dimensional projections point 3 seems to be close to points 1 and 2, but it is in fact quite distant, while points 1 and 2 remain close on any 2 dimensional projection. Point 3 is thus called a false neighbour in 2 dimensions; it is measured as being a short distance away in 2 dimensions, but upon embedding in 3 dimensions it is no longer a neighbour.

A false nearest neighbour in m -dimensions is thus a point that is the nearest neighbour in m dimensions, but is much farther in $m+1$ dimensions. Following the method of Kennel et al., (1992) a constant must be chosen, and we declare $n(\mathbf{x}(t))$ to be a false nearest neighbour if

$$\left[\frac{d_{m+1}^2(\mathbf{x}(t), n(\mathbf{x}(t))) - d_m^2(\mathbf{x}(t), n(\mathbf{x}(t)))}{d_m^2(\mathbf{x}(t), n(\mathbf{x}(t)))} \right]^{1/2} > R_{Tol}$$

Figure 2. (top) The embedding of a sine wave in 2 dimensions, the points are shaded to show their position in the plane. (bottom) The same embedding projected into 1 dimension, with the points shaded with their 2-dimensional positions to show the change in neighbours.

Figure 3. Three 2-dimensional projections of the 3-dimensional embedding of the x co-ordinate of the Lorenz attractor; the shading of points is by position in the 3-dimensional embedding. The top-left image is the most accurate, as few neighbouring points contrast in shade.

When this criterion is tested for every point, we can find the percentage of points which have false nearest neighbours (False Nearest Neighbour Percentage, FNNP); this measure of the fit of the embedding dimension should drop quickly, becoming 0 for any dimension greater or equal to the minimum embedding dimension; due however to the presence of noise this goal is not always attainable (Kennel et al., 1992).

A problem can arise when the nearest neighbour is in fact not “near” the reference point – any nearest neighbour that is more than $1/R_{Tol}$ of the diameter of the attractor away from the reference point cannot be labeled a false neighbour by the above criterion. The average distance between points will increase as we increase dimension, so this problem can dominate higher dimensions. To prevent this effect, a second criterion is also tested, to determine whether the point is too distant to be counted as a neighbour in the first test. Given a measure of the size (dispersion) of the attractor, d_A , we can set a constant A_{Tol} and consider a neighbour false if

$$\frac{d_{m+1}(\mathbf{x}(t), n(\mathbf{x}(t)))}{d_A} > A_{Tol}$$

The measure of the attractor employed by Kennel et al. (1992) was the variance of the distances from the center of the attractor, which we will follow, but they note that other choices of d_A did not alter their results. We can then use the two criteria jointly to determine false nearest neighbours, and any point that meets either criterion will be labeled as having a false nearest neighbour.

Sampling and delay

An important point is the selection of a delay; the delay should be large enough to allow the dynamics to be changing, yet not so long that the points are non-related. The

selection of the delay is not a trivial matter – although any delay will work in an ideal mathematical model (Hegger et al., 1998, Schreiber, 1998) it is important to select an appropriate delay in noisy, finite data. Hegger et al. (1998) also points out that in some systems where there are different periods within a signal there is no right embedding delay, and a choice of delay determines which feature is emphasized.

Commonly chosen delays are the first zero of the autocorrelation function, the point at which the autocorrelation drops below $1/e$, or the first inflection point of the autocorrelation function or a delay based on the decay of mutual information, as reviewed by Rosenstein et al. (1994). The reasoning for these choices is a balance between redundancy and irrelevance – the introduction of a data point that is too close temporally will provide little new information, and is redundant, while if the delay is too long the data points may not be causally related, and the embedding will not represent the underlying dynamics (Casdagli et al., 1991). The autocorrelation function's decay as a function of τ is a measure of the decay of the linear relationship, and estimates the time required for the linear correlations to become independent. The use of the first minimum of the mutual information is a better option in some ways, as it is a general rather than linear relationship that is used, but would require four or five orders of magnitude more calculation than an autocorrelation (Rosenstein et al., 1994). For the sake of computation speed and simplicity we will use the point at which the autocorrelation function drops below $1/e$ as a yardstick for our delay, using the average of the values for the different electrodes.

Due to the relationship between successive points in the original time series we find that if we search for neighbours among the embedded points they are disproportionately

found to be temporally proximate; the neighbour of $\mathbf{x}(t_n)$ would often be $\mathbf{x}(t_{n+1})$ or $\mathbf{x}(t_{n-1})$. In order to avoid comparing points which are related through temporal proximity rather than through similar underlying dynamics we wish to select as neighbours points that are temporally removed. This same issue was addressed by Arnhold et al. (1999) by use of a minimum time, τ_{Theiler} that must separate the sample point from its neighbours. To this end, we select a fraction of the points to use as a set on which to perform our neighbour tests and subsequent analyses; if every second point was used as the index for a point, we would have a sampling (s) of 1, and with a sampling of 9, every 10th embedded point is used. We can still embed all the points, and then we can have many interleaved sets of points, each set being sampled from our attractor with points in the subset less apt to be found to be dependent in the temporal sense. We can call these subsets of our embedded points *streams*, and then with a sampling s we have s streams, the n^{th} stream being

$$X_{(n)} = \{\mathbf{x}(t_i) \mid i = n + js, j \in \mathbb{N}\}$$

This approach reduces the amount of data available for the comparison, but has the advantage of being much faster to calculate than other methods. To determine how far apart sampled points should be we embed the time series, select a number of random points, and increase the width of our sampling (s) until all the selected points have as nearest neighbours points more temporally removed than s . We will be searching for neighbours among the points of the same stream, and as the points of the stream are sufficiently spaced to be somewhat independent a neighbour is unlikely to be related by reason of temporal proximity. Arnhold et al. (1999) also end up applying a similar method as, instead of embedding every point, they have simply decimated the time series by the delay τ .

Nearest neighbour distance

Once embedded in an appropriate dimension m , with a sampling s and delay τ we have s streams of points that form orbits in phase space. Each point represents the state at a point in time, which is recorded by a time index. For every point in X we find k nearest neighbours, and the average distance from the point to its k nearest neighbours, which is the true neighbour distance (tnd) at time t_n and for k neighbours.

$$tnd(X, t_n, k) = \frac{\sum_{i=1}^k d(\mathbf{x}(t_n), n(\mathbf{x}(t_n), i))}{k}$$

Referred neighbour distance

To compare two signals we now find the referred neighbour distance (rnd) for each point in X . If we wish to refer from the embedded signal Y , for each point $\mathbf{x}(t_n)$ we find the corresponding reference point $\mathbf{y}(t_n)$. The k neighbours of $\mathbf{y}(t_n)$ are found, let them be $\mathbf{y}(t_1), \mathbf{y}(t_2), \dots, \mathbf{y}(t_k)$, and their time indices are recorded ($t_1, t_2, \dots, t_k$). We find the points in X sharing those time indices, ($\mathbf{x}(t_1), \mathbf{x}(t_2), \dots, \mathbf{x}(t_k)$), and much like our notation for neighbours we will use $r(\mathbf{x}(t_n), i, Y)$ to indicate the i^{th} referred neighbour of $\mathbf{x}(t_n)$ from Y . These are the referred neighbours of $\mathbf{x}(t_n)$ from Y , and the referred neighbour distance is the average of the distances from $\mathbf{x}(t_n)$ to the referred neighbours, or

$$rnd(X, t_n, k, Y) = \frac{\sum_{i=1}^k d(\mathbf{x}(t_n), r(\mathbf{x}(t_n), i, Y))}{k}$$

Figure 4: Two time series shown, with the steps involved in finding the true neighbour distance (tnd) and referred neighbour distance (rnd). $x(t)$ is our sample point, $n(x(t), 1)$ nearest neighbour, the distance between them in phase space is the tnd . $y(t)$ is the matching time index for our sample point, $n(y(t), 1)$ is the neighbour of the reference point, and $r(x(t), 1, Y)$ is our referred neighbour from Y , with the distance in phase space from sample point to the referred neighbour being the rnd .

Figure 5: This shows the same two signals as Figure 4, but with a shift of Δt being applied; as a result of the shift we see that different points result for the mutual and referred points, and thus the divination coefficient changes. The Δt associated with the maximum of the divination function (Dv_{max}) is an estimate of the lag between signals.

where Y is the referred set of embedded points, k is the number of neighbours used, and n is the time index. The divination coefficient is then simply the average of the ratio of the referred neighbour distance to the true neighbour distance.

$$Dv(X, Y) = \frac{\sum \frac{tnd(X, t_n, j)}{rnd(X, t_n, j, Y)}}{N}$$

Divination function

Much like the cross-correlation function, the divination coefficient can be calculated comparing the two signals with an artificial delay of length Δt , so that the point $\mathbf{x}(t_n)$ is paired with the point $\mathbf{y}(t_{n+\Delta t})$; the true neighbour distance calculation remains the same, but the neighbours of $\mathbf{y}(t_{n+\Delta t})$ are found instead, and we subtract Δt from the indices before referring back to X . We thus define the *rnd* equation to include a shift Δt , so $rnd(X, t_n, j, Y_{\Delta t})$ is the referred neighbour with Y shifted Δt .

$$Dv(X, Y; \Delta t) = \frac{\sum \frac{tnd(X, t_n, j)}{rnd(X, t_n, j, Y_{\Delta t})}}{N}$$

If we find the maximum value of the divination function ($Dv_{\max}$) we have a measure of the lag in synchrony between the two signals, in the same way that maximum cross-correlation is used to determine the lag between two signals (Mormann et al., 2003; Meeren et al, 2002).

Implementation

Dimension program

A program was written using LabVIEW 5.1 to determine an appropriate embedding dimension for time series data using the false nearest neighbour method of Kennel et al. (1992, see *Embedding dimension and neighbours*). The completed program was applied to test data to ensure that they were producing results consistent with the known properties of the test data before testing was done on the biological data. The program for assessing the dimension of data was tested on sine waves, the Lorenz attractor, the Rossler attractor coupled and uncoupled with the Lorenz attractor, seizure and non-seizure activity, as well as noise contaminated versions of sine waves and the Lorenz attractor, in order to test the accuracy in noise contaminated signals.

Divination program

A program was written using LabVIEW 5.1 to calculate the $Dv_{\max}$ of time series data (and the divination coefficient), using the method outlined in the sections *Nearest neighbour distance*, *Referred neighbour distance* and *Divination function*. Since the divination function is a novel measure, in order to assess whether the program is functioning properly it was tested against time series that had predictable divination functions. The program was tested on uniform noise, sine waves of same frequency, in and out of phase, sine waves of differing frequency, in and out of phase, the Lorenz attractor, systems of coupled attractors and EEG data with and without a time shift applied.

Biological data

The divination function program was applied to recorded electrical activity from depth electrodes implanted bilaterally in the hippocampi and amygdalae of rats. The rats were of two strains, Fast and Slow, describing the ease of kindling and resistance to kindling in these strains respectively. The recordings are from 11 rats, 5 from the Fast strain and 6 from the Slow strain. Recordings were made from depth electrodes located in the left and right hippocampus and left and right amygdalae (4 electrodes), and were amplified using a Grass amplifier (Grass Inc.; West Warwick, USA) with a built in lowpass filter ($f_{1/2}$ = 75Hz) and stored on VHS tape. The data was subsequently digitised at 1000Hz, and digitally bandpass-filtered between .5 and 200 Hz.

Spontaneous recordings

The 30-second segments of activity were selected from 30-minute recordings of each rat pre- and post-kindling, and were selected due to the similar behavioural notes made about the rat at the time of the segments. For each rat a segment was selected with the rat immobile, and a segment in which the rat was sniffing, for both the pre- and post-kindling recordings. The immobile and sniffing conditions are treated separately as the behaviours represent different states of brain activity. Dv_{max} was calculated on 15 2000 ms non-overlapping windows, with a sampling (s) of 5 for all signals in order to allow for comparisons. The resulting Dv_{max} values were analysed using SPSS 11.0 using a mixed design general linear model approach.

Stimulated after discharges

The divination function program was also applied to the recordings of the first elicited seizure during the afterdischarge threshold testing. The afterdischarge is a large amplitude seizure following electrical stimulation during the kindling process, and this is thus the first elicited seizure. These recordings vary in length with the length of the seizure, and analysis of these signals was done on 3000 ms overlapping windows, with a 200 ms delay between consecutive windows. $Dv_{\max}$ was recorded for each window, as well as the associated Δt .

Results

As the goal of this thesis was the development of a method for measuring general synchronies, much effort was devoted to the development of the programs, restructuring the programs to improve speed, and testing of the programs on data types with predictable or known characteristics. This defines the initial part of the results section, and will be followed by the results of testing the programs on biological data.

Properties of the divination coefficient

As the properties of the divination coefficient have not previously been explored some basic results will now be given. I will focus on the four most important results, which are:

1. The divination coefficient lies between 0 and 1.
2. The divination coefficient can be non-symmetric.
3. For pairs of periodic signals the divination coefficient is unaffected by their relative phases
4. For the divination function there are two measures of a leading/lagging signal, and they can be in disagreement.

Value of the divination coefficient

The divination coefficient quantifies the ability to predict the states of one attractor from the states of a second attractor, and can assume values in the range $(0,1]$, with a value of 0 being a low degree of predictability and a value of 1 being a high degree of predictability. The value has an upper limit of 1 because it is calculated by dividing the nearest-neighbour distance by the referred-neighbour distance, and as the nearest-

neighbour distance is the smallest distance from the reference point to any other point, it must therefore be smaller than the referred-neighbour distance. The divination coefficient is non-negative as it is always the quotient of two positive non-zero numbers. The minimum possible value occurs when every referred neighbour is the furthest point in the embedding from the reference point. Therefore the minimum value is in the range of the nearest neighbour distance divided by the diameter of the embedding, and the divination coefficient can only ever be arbitrarily close to 0.

Non-symmetry

Divination coefficients are measured between two time series much like correlation coefficients, but are different in that a divination coefficient is non-symmetric, i.e. the value of $Dv(A,B)$ is not necessarily the same as $Dv(B,A)$, while the correlation coefficient must always be symmetric. Some time series have greater predictive power in one direction (B predicting A) than in the other (A predicting B), and this is seen most easily in the case of periodic functions with different frequencies. An example is two sine waves with frequencies of f and $12f$. If we think of this system as representing the hands of an analog clock we can see that the slow signal provides us with much more information about the fast signal than vice versa; a clock missing the minute hand still reliably tells time with some degree of accuracy as the hour hand provides us with all the information we need, with one position for each position of the minute hand. A clock missing the hour hand only indicates that the time is (for example) quarter past-something, indicating that the hour hand would be at one of 12 positions. This can be generalised to the case of frequencies f and nf , and the divination coefficient reliably shows these asymmetries.

Phase insensitivity

Periodic signals are common in nature, and as a result the ability of the divination coefficient to deal with them is important. If we examine a simple case, that of two sine waves of the same frequency, we find that when they are in phase both the divination and correlation coefficients are 1; the highest possible value. For the correlation coefficient, shifting the signals by one half period results in a value of -1 (also a strong relationship, but negative). In fact, by choosing phase offsets of the appropriate amount we can generate any correlation coefficient between -1 and 1 . The divination coefficient on the other hand is equal to 1 for all phase shifts. This insensitivity to the phase of a signal can be a strength, as it can potentially allow us to separate the contribution of the periodic and non-periodic contributions of signals to the divination coefficient.

The maximum cross-correlation has been used to determine the lag between two signals (Meeren et al, 2002; Mormann et al., 2003), arguing that if two signals are coupled in some manner that the offset time associated with the maximum of the cross-correlation function may be a measure of the time required for a response, and that the earlier signal may drive the later signal. In signals with a periodic or quasi-periodic structure this may simply result in determining the phase difference, and as the coupling need not be a linear relationship the maximum cross-correlation fails to detect many forms of synchrony.

If we record $Dv_{\max}$, the offset then represents the time difference that allows the best prediction from one signal to another, and as it is insensitive to the phase of the signal we are not simply detecting a phase offset, but a true improvement in prediction over this time lag. In situations for which a delay might be expected, and when theory suggests

that the signals are coupled, this lag may be a good measure of the delay between the signals.

Lag and directionality

The divination function between two signals is a measure of how well one can predict from one embedding to another at various delays. If we are choosing to use the $Dv_{\max}$ as a measure of the strength of the relationship and the associated Δt as a measure of the delay occurring between the signals it is important to be aware of possible interpretation difficulties. One such problem arises in that we can have non-symmetry in the delays at which the $Dv_{\max}$ occur and in the strength of the $Dv_{\max}$ observed.

We know that non-symmetric $Dv_{\max}$ can be observed between signals with different frequencies for example, and that interpreting a difference in the $Dv_{\max}$ as indicating a causal relationship is thus not justified, though if a driving function exists and is driving several functions the expectation is that the predictions would be better in the direction of the driving force. Similarly, there exist situations in which the delay associated with $Dv_{\max}$ is due simply to random fluctuations in Dv associated with the various delays and does not represent a delay in reaction or transmission. This is most likely to happen when the divination function has no clear maximum, due to low levels of synchrony between the signals, such as in the case of uncoupled signals.

Truly confusing situations can thus arise, in which the $Dv_{\max}$ is asymmetric, indicating a possible driving force, and the delays associated with the $Dv_{\max}$ could both be positive, both be negative or simply be of different magnitudes. It is difficult to conceive of actual situations in which each signal had the best divination value when compared with (for example) an earlier version of its reference signal, yet this can arise, especially when we

deal with noise or independent signals. As a general rule we will treat the delay as being a true measure of an improved synchrony at that delay only when the divination function is significantly elevated. This requires us to be able to judge whether a given $Dv_{\max}$ is significant, which is an issue we will take up in the Discussion.

A slight complication can arise when searching for the $Dv_{\max}$, due to the presence of false peaks in the divination function that arise as a function of the delay and sampling used. Due to the way in which points are formed (by taking regularly spaced data points to form the coordinates) an offset equal to the delay on an unsampled time series will often result in false peaks, as the offset reference point shares most of its coordinates with the reference point at 0 offset. In the event that the attractor is sampled, the same error occurs if the sampling is done at a factor of the delay, or to a lesser extent if the lowest common multiple of the sampling and delay is less than $(m-1)\tau$. To avoid this we can select the delay in such a way that we minimize this problem, by selecting a prime number for our delay and our sampling as well if necessary, which is the precaution taken in the programs (using the nearest prime to the delay, and a different prime for the sampling of the attractor).

Choice of parameters and optimization

To calculate the divination function on data several decisions must be made based on the type of data to be analysed and the desire to improve computation speed. In early runs sub-optimal choices of parameters resulted in a large number of operations, causing the computation time to be over 48 hours for a signal of 60000 data points (60 seconds) from four electrodes. This was unacceptable, as the total computation time given the

number of signals to be analysed could have exceeded a year. As a result of much experimentation and theory the choices of parameter and methodology outlined below were made based on the data type and the desired speed and accuracy.

The first parameter to be considered is the delay. Both the divination program and the dimension program were written to evaluate an appropriate delay for the embedding, the τ for which the autocorrelation function dropped below $1/e$ was used as a base for this value, selecting the nearest prime above this value. These values of τ provided reasonable spread of the attractor from the diagonal by visual inspection, and are largely in the range of 40-60 ms, with a few instances as low as 29 ms or as high as 101 ms.

The sampling was determined automatically for the divination program as well, using the method outlined in *Sampling and delay*, and selecting the first prime greater than the sampling value calculated. A check was performed to ensure that the same prime was not used for both the delay (τ) and the sampling (s), and if necessary the sampling was increased to the next prime. The sampling ranged from a low of 7 to a high of 23, with most of the signals lying in the range of 11-17.

The program to calculate the minimum embedding dimension requires 3 parameters, the A_{Tol} and R_{Tol} described in *Embedding dimension and neighbours*, and a length for the signal to be analysed. The A_{Tol} and R_{Tol} were selected to be 2 and 15 respectively, based on the accuracy they had in evaluating the dimensions of our toy data sets, and the values reported in Kennel et al. (1992). The length of the data set was 1500 points, as this was enough to give several orbits of any of the attractors we were examining, and still remained both fast (<10 minutes for 1500 pts on 4 channels) and accurate, reporting the

values in the same range as those calculated using 6000 points, which had taken ~ 2 hours per 6000 point sample.

For the divination function program a window length was needed, as well as the decision of whether to have overlapping or non-overlapping windows; if overlapping windows were chosen, the time between the start of each window had to be decided on. Different window lengths were used for the stimulated and spontaneously recorded data, as they have different values of sampling and delay.

For the stimulated seizure data it was decided to use overlapping windows, in order to have a smooth picture of how the synchronies were changing with respect to time. The time between windows was set to be 200 ms, as this is a period short enough to remain relatively stationary, while still long enough to allow quick computation times. Given the embedding dimension of 6 and the sampling of ~ 15 , this results in only moving the windows by ~ 13 points. For the analysis of the spontaneous data (both sniffing and immobile conditions) we wanted to have many independent estimates of the resting levels of the divination coefficient, so the windows were chosen to be non-overlapping.

The length of the windows is chosen based on several factors, among them the dynamics of the system (the characteristic time over which a signal remains quasi-stationary), the number of orbits within the window if the signal has an attractor, the desired accuracy and the desire for speedy computation. As the window is lengthened the number of orbits increase, making closer nearest neighbours more likely. Closer nearest neighbours increase accuracy of the prediction, but if the window is too long in a system with shifting dynamics the attractor may shift during the window, and the result

would be less accurate. Making the window shorter decreases computation time, but potentially also affects accuracy as the neighbour matches become more random.

The choice was made to use windows 3000 ms long for the stimulated seizures, as the brain is known to be relatively stationary over this time, and it provides enough points for several orbits of the attractor and thus is believed to be accurate. A shorter 2000ms window was selected for the spontaneous recordings as the signals did not have the seizure related spiking at 2-3 Hz, and thus has more orbits in a shorter window.

Testing of the decay of the divination function as the lag between signals increased showed a tendency for the value to decrease fairly rapidly in all signals with a form of general synchrony. The biological data was tested initially with lags up to 400 ms in either direction. It was observed that over 95% of the time the peak occurred within a lag of ± 25 ms if the peak was substantially above noise, and when the peak failed to lie in this range for the test data it was generally associated with a low value of the divination function, often within the range of a comparison with noise or a surrogate shift (surrogate shifts are a form of significance testing, and will be discussed later). On this basis the range of the lags was reduced to 25 ms in either direction, reducing the number of calculations by a factor of 16, and resulting in a significant reduction in computation time.

To further reduce the amount of data processed the program was modified to use only one of the s streams that were in the embedding. With s streams one gets s estimates of the divination function, and it was noted that these values always lay in a very small range. This is hardly surprising, given that s is selected to be the smallest number for which the nearest neighbour is not temporally proximate. As a result, each of the s

streams is nearly identical to the next stream, and produces very similar values for the divination coefficient. Since s is in the range of 7-23 for our data, this results in a reduction in the number of operations by a factor of 7-23. As a result, we have approximately 60 embedded points per second, which is close to the 34.6 embedded points per second used by Arnhold et al. (1999).

Though the choices to use only one stream and to search for the peak over a more restricted range may have a small effect on the accuracy of the calculation, they are justified in that the divination coefficient did not change more than 5% from the values using s streams and a wider range. The computation time for 60 seconds of data was reduced from over 50 hours to ~ 30 minutes, reducing the total computation time for a single pass of all the biological data from roughly 1 year of computer time to only ~ 90 hours.

Determining dimensionality of data sets

A provisional decision was made to use the point at which the false nearest neighbour percentage (FNNP) fell below 1% (following Kennel et al., 1992) as being a good indicator of the minimum embedding dimension of the data. The FNNP was calculated for dimensions 1 through 9, as all of the signals were expected to be embeddable in this range of dimensions.

Figure 6 shows the relationship between FNNP and embedding dimension for several simple test systems. Overall the FNNP using a 1% threshold accurately predicts the embedding dimension we would expect based on theoretical properties of these signals. For instance, the FNNP for sine waves of various frequencies fell below 1% for dimensions ≥ 2 . This is consistent with the intrinsic dimensionality of sine waves (2

Figure 6. Graph of FNNP vs. Embedding Dimension for simple systems. Sine waves and the Rossler attractor are determined to require 2 dimensions to embed, while the Lorenz attractor and Rossler coupled to Lorenz require 3 dimensions, and noise cannot be properly embedded in 9 or fewer dimensions. All the graphs of FNNP are plotted with the $\log_{10}$ of FNNP; as a result, the values cannot be displayed for an FNNP of 0.

dimensions). At another extreme, uniform noise, which is theoretically of infinite dimension never fell below our 1% FNNP threshold. The other signals shown in Figure 6 matched their theoretical predictions.

The Lorenz attractor, theoretically of fractal dimension 2.07 (Lewkowicz et al., 2002) was found to have a minimum embedding dimension of 3. The Rossler attractor, which has the fractal dimension 2.13 (Lewkowicz et al., 2002), was found to have a minimum embedding dimension of 3. When the Rossler attractor is coupled with the Lorenz attractor we see that the FNNP falls more slowly than the uncoupled Rossler, though still falling below 1% for a dimension ≥ 3 (Figure 6). This evidence suggests that the 1% FNNP guideline is appropriate for our data, and we will adopt it as our dimensionality criterion. The Rossler and Lorenz attractors are described in Appendix B.

Effect of noise on determining embedding dimension

Since signals are often recorded in the presence of noise, it is important to verify that the method used to determine the embedding dimension is accurate in the presence of modest amounts of noise. Two types of signal were contaminated with uniformly distributed noise, a simple periodic signal (sine wave) and a more complex quasi-periodic signal (Lorenz attractor). Figure 7 shows that contamination of the sine waves with uniform noise gradually moves the curve of the FNNP upward, with the highest levels of noise resembling the outcome for noise in Figure 6. The contaminating noise level is expressed as a signal to noise ratio (SNR), with the SNR calculated as $20\log_{10}(R_A/L)$, where R_A is the standard deviation of the signal, and the noise is distributed on the range

Figure 7. FNNP vs. Embedding dimension calculated for noise contaminated sine waves. The dimension remains accurate with to within one dimension of the true theoretical value of 2 for $\text{SNR} \geq \sim 21$. The data is still considered embeddable with as little as a 10.75 dB SNR, and ceases to have a low dimensional attractor found with ≤ 7.25 dB SNR, behaving like noise at this level of contamination.

Figure 8. Graph of FNNP vs. noise contaminated Lorenz attractor. The predicted minimum embedding dimension remains accurate within one dimension of the true embedding dimension of 3 for $\text{SNR} \geq 19$ dB.

Figure 7.

Figure 8.

(L,-L). The minimum embedding dimension rose from 2 dimensions to 3 dimensions for the sine waves up to a signal to noise ratio (SNR) of about 20 dB, and rising to 4 dimensions for sine waves with ~10dB SNR.

Figure 8 shows that the Lorenz attractor similarly suffered an increase in the calculated minimum embedding dimension as theory would predict, with the minimum embedding dimension remaining 3 dimensions with up to ~30dB SNR, and only rising above 4 dimensions with ≤ 17 dB SNR. These results show that noise causes an increase in the FNNP, but that without the signal being below ~20dB SNR there is little effect in either example.

Dimension and biological data

The results of tests on spontaneous activity recordings of Slow and Fast strain rats before and after kindling are shown separately for the case of immobile rats and for rats engaged in sniffing type activities. Figure 9 shows that the FNNP drops below 1% for the immobile rats at a dimension of 5 (average was 4.59) for the kindled rats, and a dimension of 6 pre-kindling (average 5.75); this holds true in Figure 10 for the rats while sniffing, which also drop below the 1% FNNP threshold at 5 dimensions for the kindled rats (average 4.59), and at 6 dimensions pre-kindling (average 5.43). The Slow and Fast strains of rat were not distinguishable by the FNNP ($F(1,32)=1.537$, $p=.224$ for sniffing condition, $F(1,32)=1.007$, $p=.323$ for immobile condition). The FNNP were significantly different for the pre- and post-kindling groups with a $p<.001$ ($F(1,32)=15.849$) for the immobile condition and a $p<.001$ ($F(1,32)=15.110$) for the sniffing condition.

The embedding dimension was also calculated for the case of stimulated activity, with the data used being the first after discharge stimulated from a hippocampal stimulus. The

Figure 9. Graph of FNNP vs. Embedding Dimension for the spontaneous activity of the immobile rats, both pre- and post-kindling. The FNNP drops below 1% for the immobile rats at an average dimension of 4.59 for the kindled rats, and an average dimension of 5.75 pre-kindling, a significant difference; $p < .001$ ($F(1,32)=15.849$).

Figure 10. Graph of FNNP vs. Embedding Dimension for the spontaneous activity of the sniffing rats, both pre- and post-kindling. The FNNP drops below 1% for the sniffing rats at an average dimension of 4.59 for the kindled rats, and an average dimension of 5.43 pre-kindling, a significant difference; $p < .001$ ($F(1,32)=15.110$).

Figure 9.

Figure 10.

FNNP for each of the stimulated hippocampus (4.82), unstimulated hippocampus (5.28), ipsilateral (4.91), and contralateral amygdala (4.91) dropped below 1% at an average dimension of 4.98 (Figure 11).

In the afterdischarge elicited by amygdaloid stimulation similar results were seen, with the stimulated amygdala dropping below 1% FNNP at 3.91 dimensions on average, and the contralateral hippocampus (4.73), ipsilateral hippocampus (4.55) and unstimulated amygdala (4.91) areas falling below 1% FNNP at about 5 dimensions, with the average being 4.53 (Figure 12). An ANOVA was performed on the data from both the hippocampal and amygdaloid stimulated seizures to test the effect of electrode location and strain of rat, and no effect of electrode location or rat strain was found in either the hippocampal stimulation ($F(3,40)=0.45$ for electrode location, $F(1,40)=0.547$ for strain), or the amygdaloid stimulation ($F(3,40)=1.883$ for electrode location, $F(1,40)=0.565$ for strain).

Figure 13 shows how within a subject the embedding dimensions can vary between the monitored regions, with embedding dimensions of 2, 4, 4 and 5 for the stimulated hippocampus, unstimulated hippocampus, ipsilateral amygdala and contralateral amygdala respectively. These variations are common, as the definition of an after discharge is the presence of a discharge in one of more regions lasting 3.5 seconds or longer (Racine et al., 1999) and could thus vary in duration, intensity, generalization and which in which region(s) it presents itself. Figure 14 shows the range of dimensions for all rats as measured for the stimulated hippocampus in the hippocampal stimulation condition, and the point at which the FNNP falls below 1% varies from the 2nd to 7th dimensions. Based on the above evidence the choice was made to perform the divination

analyses at 6 dimensions, which is consistent with that selected for the spontaneous recordings.

$Dv_{\max}$ and noise

The divination coefficient for two time series varies with the embedding dimension and the number of embedded points. For time series that are only predictive by chance the divination coefficient is the average of the ratio of the nearest-neighbour distance and the average distance, for a random selection of points on the attractor. The expected value of the divination coefficient between independent signals thus varies in a systematic way, and this can be shown most easily with two time series consisting of uniform noise.

The average neighbour distance for uniform noise will obviously decrease as we populate a given space with an increasing number of points. Given a cube with 100 points randomly scattered through it they are bound to be further apart on average than 1000 points in the same cube would be, and this holds in any dimension of hypercube; thus within a given dimension of embedding, the more points we use, the lower the average neighbour distance. As we increase in dimension we can expect to see the average neighbour distance increase, however, as we must spread the finite number of points through a larger space – if 100 points are scattered at random on a square, and then scattered at random into a cube we will find that they are more distant in the cube, though this is very difficult to state in a precise manner mathematically. We can see that logically, the N points can be approximately spread out equally in all D dimensions provided the number of points is large ($N \gg D$), and that the average neighbour distance

Figure 11. Graph of FNNP vs. Embedding Dimension for seizure resulting from hippocampal stimulation. The average FNNP is included as no differences were found for the FNNP between regions.

Figure 12. Graph of FNNP vs. Embedding Dimension for seizure resulting from amygdaloid stimulation. The average FNNP are shown as no differences were found for the FNNP between regions.

Figure 11.

Figure 12.

Figure 13. Graph of FNNP vs. Embedding Dimension in a hippocampally stimulated rat, showing the variation in dimension between the regions.

Figure 14. Graph of FNNP vs. Embedding Dimension in the stimulated hippocampus in all hippocampally stimulated rats, showing the variation in dimension within a given recording site. The solid markers and the cross are Fast rats, while the hollow markers, X and star are the Slow rats.

should be roughly proportional to the D^{th} root of the number of points, or $D_n \sim (1/N)^{1/D}$.

This is stated in more precise terms in Bhattacharyya & Chakrabart (2003).

The average distance between two points chosen at random in a hypercube (and thus in embedded uniform noise) however only varies with the dimension of the hypercube, this distance increasing as the dimension is increased. The divination coefficient predicting uniform noise from any other system is a ratio of these distances, with the numerator decreasing as the number of points (N) increase, and as the dimension increases the numerator increases substantially, with a smaller increase in the denominator. This results in the divination coefficient decreasing as N increases, and increasing when m (embedding dimension) increases, and this relationship is shown in Figure 15. This gives us a conservative estimate of the divination coefficient for randomly associated signals, which I will use in judging the significance of the Dv_{max} . I will examine alternatives to this in the discussion.

Verification of the divination program

The program to calculate Dv_{max} cannot be compared against known values (as no previous use of this method exists), but can be tested against theoretical values in some cases, or for the trend it should observe in other cases. The program was tested on uniform noise, sine waves of same frequency, in and out of phase, sine waves of differing frequency, in and out of phase, the Lorenz attractor, systems of coupled attractors and EEG data with and without a time shift applied. Where expected values were known for these systems the values were tested using independent sample t-tests, with all results presented in Table 1.

The calculated $Dv_{\max}$ for noise were in line with the expected values, and clearly show the increase in the $Dv_{\max}$ when dimension is increased, and the decrease in value when the number of points is increased (Figure 15).

Same frequency sine waves both in and out of phase have a $Dv_{\max}$ of 1.0 (Table 1) and sine waves of different frequencies predict better from the slower wave to the faster when they are multiples, and are unaffected by phase differences (Table 1).

The $Dv_{\max}$ between the coordinates of the Lorenz attractor were significantly higher than noise, and a one-way coupling of the X value for the Rossler attractor with the X value from the Lorenz (with the Lorenz driving the Rossler) resulted in a decrease in $Dv_{\max}$ between the X and Y of the Rossler attractor, while the $Dv_{\max}$ predicting the Rossler attractor from the Lorenz attractor increased, despite no change occurring in the opposite relationship (Table 1).

In Figure 16 we have an EEG from a “sniffing” rat with the second half of the signal rotated by 2500 ms to create a shift surrogate. There is a clear drop in the value of the $Dv_{\max}$ at the time that the shift happens, and the average $Dv_{\max}$ in this case drops from .070524 to .053750 (Table 1). The same set of data with the order of all points randomly permuted also shows significantly lower $Dv_{\max}$ compared to the non-permuted sample, with an average $Dv_{\max}$ of .04056, while a sample of uniform noise has an average of .041333 (Table 1). The $Dv_{\max}$ was unaffected by any scaling or translation of the time series, as was predicted by theory (not shown, all values remained exactly equal).

Figure 15. Effect of number of points and dimension on the divination coefficient. As the number of points increase the divination coefficient decreases, and with an increase in embedding dimension the divination coefficient increases.

Figure 16. A graph of the $Dv_{\max}$ vs. time for a system where the indices of the time series values for one of the signals has been shifted by 2500 ms to form a surrogate (i.e. at times greater than 7 s). The values of $Dv_{\max}$ show a significant drop after the temporal relationships have been altered in this way ($t=p<.005$). Also visible is the $Dv_{\max}$ for the same signals with one set of vectors permuted, and the divination coefficient for a comparison of the first signal against an equal length of uniform noise.

Figure 15.

Figure 16.

Test systems	Divination Coefficient	Theoretical Value ^a	p value ^b
(1) Sine waves, same frequency, same phase	1.0000	1.0000	n/a
(2) Sine waves, same frequency, different phases	1.0000	1.0000	n/a
(3) Sine waves, 1:2 frequency, same phase	0.0050	n/a	n/a
(4) Sine waves, 1:2 frequency, different phases	0.0050	= (3)	>.05
(5) Sine waves, 2:1 frequency, same phase	0.2114	> (3)	***
(6) Sine waves, 2:1 frequency, different phases	0.2099	= (5)	>.05
(7) Sine waves, 1:3 frequency, same phase	0.0040	n/a	n/a
(8) Sine waves, 1:3 frequency, different phases	0.0040	= (7)	>.5
(9) Sine waves, 3:1 frequency, same phase	0.0950	>(7)	***
(10) Sine waves, 3:1 frequency, different phases	0.0911	= (9)	>.05
(11) Sine waves, 2:3 frequency, same phase	0.0050	n/a	n/a
(12) Sine waves, 2:3 frequency, different phases	0.0050	= (11)	>.5
(13) Sine waves, 3:2 frequency, same phase	0.0030	n/a	n/a
(14) Sine waves, 3:2 frequency, different phases	0.0030	= (13)	>.5
(15) Lorenz attractor X -Y	0.6907	> noise	***
(16) Rossler X - Y, pre coupling to Lorenz	0.8309	n/a	n/a
(17) Rossler X - Y, post coupling to Lorenz	0.4964	< (16)	***
(18) Rossler X - Lorenz X, pre coupling	0.0416	n/a	n/a
(19) Rossler X - Lorenz X, post coupling	0.0904	> (18)	***
(20) Lorenz X - Rossler X, pre coupling	0.1111	n/a	n/a
(21) Lorenz X - Rossler X, post coupling	0.1217	= (20)	>.05
(22) Depth electrode recordings, sniffing	0.0761	n/a	n/a
(23) Depth electrode recording, surrogate shifted	0.0538	< (22)	*
(24) Depth electrode recording, permuted vectors	0.0389	< (22)	***

^a Theoretical value of the divination coefficient; in many cases the only restriction on the value is that it be larger, smaller or equal to the value for the test system indicated in parentheses. The symbols “>” and “<” indicate that a one-tailed independent sample t-test was performed between the value calculated and the value calculated for the system indicated in the parentheses while the “=” symbol indicates that a two-tailed t-test was performed.

^b Result of the t-test indicated in the previous column.

*p<.05, ** p<.01, *** p<.001

Table 1. Calculated values of the divination coefficient, $Dv(X,Y)$, and results of the t-tests performed on the test systems. In each case the first system described is being predicted with the predictor listed second, i.e. “sine waves 3:1 frequency” is the prediction of the frequency $3f$ sine wave from the frequency $1f$ sine wave.

Spontaneous recorded activity

Electrical activity in the hippocampi and amygdalae of Fast and Slow strain rats was classified as “immobile” or “sniffing” based on notes of the exhibited behaviour during the recording; these recordings included periods prior to any kindling and after full kindling had been achieved. A mixed design ANOVA was performed, with kindling as the within subject variable and strain of rat the between groups variable for each of the conditions, sniffing and immobile. In both the immobile and the sniffing conditions there was a significant effect of the kindling*strain interaction, with both having $p < .05$ ($p = .036$ for sniffing, $p = .026$ for immobile). For both conditions a post-hoc test of differences showed a difference between strains pre-kindling, and failed to detect any difference post kindling; in the immobile condition the Slow rats had significantly lower Dv_{max} than the Fast rats did pre-kindling ($p = .037$) for (Figure 17), while for the sniffing condition the reverse was true, with the Slow rats having higher Dv_{max} than the Fast rats pre-kindling ($p = .021$) (Figure 18). In addition, in the immobile condition the Dv_{max} of the Slow rats was increased following kindling ($p = .013$).

Stimulated after discharges

Dv_{max} measured during seizure were recorded for the first after discharges elicited from both the hippocampus and the amygdala for all rats, and though it is not possible to compare directly between these seizures due to differences in the duration, extent and pattern of the synchronies, it is clear that several things occur. The Dv_{max} rises far above the resting levels in all cases; the peaks do not occur at the same time in all channel comparisons, rather it seems that the patterns of synchrony can shift between regions.

Figure 17. Graph of the $Dv_{\max}$ pre- and post-kindling for the Fast and Slow rat strains in the immobile recording; there is a significant pre-kindling difference in the $Dv_{\max}$ ($p=.037$), but no evidence of a post-kindling difference. The Slow rats showed a significant increase in $Dv_{\max}$ following kindling ($p=.013$).

Figure 18. Graph of the $Dv_{\max}$ pre- and post-kindling for the Fast and Slow rat strains in the sniffing recording; there is a significant pre-kindling difference in the $Dv_{\max}$ ($p=.021$), but no evidence of a post-kindling difference.

Figure 17.

Figure 18.

Figure 19 shows examples of $Dv_{\max}$ from 3000ms windows 200ms apart during a hippocampally stimulated seizure, and we see that the regions enter into synchrony with each other at different times. For instance, the $Dv_{\max}$ drops in Figure 19c) after 80 windows; this would correspond to a window about 16 seconds into the seizure, and if this is compared to the graphs of the electrode recordings in Figure 20, we see that this coincides with the offset of the large amplitude spiking seen at 18 seconds on the two hippocampal electrodes. As well, we see a large rise in the $Dv_{\max}$ between the left hippocampus and left amygdala over the windows 100-120, which would be associated with the windows starting from 22 seconds to 24 seconds, and shows a change in the character of the amygdaloid signal to match the signal in the hippocampus (Figure 20). Complementing the changes in $Dv_{\max}$, Figure 21 shows the record of the temporal lags between signals accounting for the $Dv_{\max}$. The record of the lags for the two hippocampal electrodes show that the lags are consistent in direction, with the direction of the lags indicating that the unstimulated hippocampus is leading by about 5ms, but the lead progressively shrinks as we approach the offset of the first spiking stage, until it is at no lag for the offset itself – in fact the lag has dropped to zero by about the 65th window, and $Dv_{\max}$ does not fall until the 80th window. It would seem that the lag disappearing between the hippocampi is predicting the offset of seizure and the drop in $Dv_{\max}$ by ~3 seconds; this pattern does not show up in all seizures, but is present in some, with the point of lag decay or reversal being shortly before the seizure offset. The lag between the left hippocampus and left amygdala seems to fluctuate slightly until the ~120th window, by which point it has settled to 0, where it remains for the duration of the signal – this is

coincidentally the point at which $Dv_{\max}$ has risen between these regions. To note all changes in the $Dv_{\max}$, lags between signals and the associated changes in the recordings goes beyond the scope of this thesis, as each seizure has a unique divination dynamic and this was an exploration of the abilities of the divination function program to detect changes in synchrony as associated with changes in brain activity, which it seems amply able to do.

Figure 19. Graphs of $Dv_{\max}$ vs. time for a seizure elicited by stimulation of the right hippocampus, the end of the stimulus is at time 0. The time axis indicates window number; 5 windows=1 s. The four circles show which regions are represented in the graph. The top circles represent the hippocampi (left = left), the bottom circles represent the amygdalae (left = left). The arrows show the direction of prediction for the solid and dashed lines. The horizontal line indicates 4.472 standard deviations above the $Dv_{\max}$ for uniform noise.

Figure 20. Graphs of voltage (μV) vs. time (ms) of the seizure from which Figures 19 and 21 are generated. The middle of the windows used for the Dv_{max} program are indicated underneath the 4th graph by the legend indicating the numbers 0 to 292. The stimulus started at $\sim$ time 0 and the offset of the stimulus is visible just after $t=2000$.

Figure 21. Graphs of the time lags associated with $Dv_{\max}$ for the stimulated seizure shown in Figure 20. The time axis indicates window number; 5 windows = 1 s. The four circles show which regions are represented in the graph. The top circles represent the hippocampi (left = left), the bottom circles represent the amygdalae (left = left). The arrows show the direction of prediction for the solid and dashed lines.

Discussion

There are essentially three types of results from this thesis; the construction of a suite of programs for the analysis of the divination function, the use of test data to discover the properties of the divination function and thus its possible interpretation and application to data, and the results of the application of the programs to biological data. I feel that the most important of these results are the creation of a program capable of detecting changes in synchrony in a variety of data types, the surprising changes in minimum embedding dimension occurring in kindling, and the presence of differences in the synchronies between regions of the brains of Fast and Slow rats before kindling. I will first discuss the results of the construction of the program itself and the exploration of the theory of divination, and then I will examine the biological results. Finally I will explain how changes to the program can be made to improve accuracy, and the discovery of a change to the method that would result in a fascinating way to improve the ability to detect changes in the divination function and to improve the statistical basis of detecting significant synchronies.

The divination function

As was hoped, the measure is sensitive enough to detect changes in synchrony during seizure, and is sensitive enough to detect differences in the resting levels of rats of different strains. The main strengths of this method are its insensitivity to the phase of periodic signals, the fact that it is a measure between two regions, and the intuitive interpretability of what is being measured.

Possibly the most surprising aspect of the divination function to come up in the exploration of its properties is its insensitivity to the phase of two periodic signals, as was shown in Table 1. This distinguishes it from other measures of relationship between signals such as cross-correlation, phase-correlation or phase coherence, all of which have been used recently in epilepsy research (Mormann et al. 2000, 2003; Netoff and Schiff, 2002). Cross-correlation is particularly poor at handling periodic signals, resulting in nearly random values as the phase of the signals are altered. The most sensitive techniques currently used to predict seizure onset are phase coherence and phase-correlation types, which have demonstrated prediction times of up to one hour (Mormann et al., 2003). The striking feature is that these measures are concerned only with the phase of the signals, the very feature ignored by the divination function while shifted, while the amplitudes are ignored by phase-correlation, and are important for identifying neighbours within the divination approach. The success of the divination function in detecting differences in synchrony suggest that there must be important information in the amplitudes of the signals that the phase-based approaches ignore, and thus the method may be a useful complement to the phase-based methods.

The result of testing the divination function on a biological signal and both a time shifted and a permuted version of the reference signal (Figure 16) uncovers an interesting way to quantify the contribution of periodicity to the divination value. A time-shifted copy of a signal (a shift surrogate) has most of the same features that the original signal had (mean, variance, spectral power), but has broken the temporal associations between specific events. When we perform this type of shift on sine waves, which are purely periodic in nature, we see no change in the value of the divination function (Table 1), but

on quasi-periodic signals we see that the value drops significantly, as in the example of sniffing data (Figure 16). The drop is not to chance levels, so some predictability exists between states of the two signals. The magnitude of the drop might then be interpreted as being the non-periodic component of the divination function.

If this is the case we should be able to remove the remainder of the divination function's value by eliminating the periodic component. This has been shown in Figure 16 as well; the permuted vectors are not time shifted, but simply randomized, this breaks both the periodic structure of the signal and the temporal associations between events, while still maintaining the same embedded points. The permuted points occupy the same hypervolume as the original embedded signal, and have the same neighbour distances, mean and variance. Permuting the vectors results in a drop to roughly noise levels, showing that all but chance associations are removed, and since presumably the temporal associations for the non-periodic components had been reduced to chance levels by the random surrogate shifts this suggests that the component left after the surrogate shifts is the periodic component.

Although it is true that a signal with periodic and non-periodic components would behave in this way, it may not be the case that this is the only way in which this could occur – it may be possible to construct a pair of signals that do not drop to noise levels when surrogate shifted, yet have no shared periodic relationship, though I have not yet managed to do so. If the conjecture holds, the comparison of surrogate shifts and randomizations of the embedded points provides us with an interesting way to attribute the portion of the divination function that is due to periodicity, and that portion which is not due to periodicity.

Though only explored briefly in the testing, the fact that $Dv_{\max}$ is a non-symmetric measure is an important strength. We showed in the testing of the Rossler-Lorenz coupled system (Table 1) that a one directional coupling between two chaotic attractors can be detected in this way. While there are instances in which a similar asymmetric divination function can be caused by periodic signals of different frequencies (such as the sine waves of Table 1) the ability to detect one directional relationships is a potentially powerful tool, and may help to deduce where the driving region is for events such as epileptic seizure, and provides us with a possible way of locating the origin of seizures. As Arnhold et al. (1999) mention, interpretation of the directionality of the divinations as being indicative of a driving and response system is only justified if there is a reason to suspect that one region drives another. As well, the interpretation of the asymmetry as indicating a causal or driving relationship does fail in situations of periodic functions of different frequencies, and presumably quasi-periodic functions of different frequencies. It should be possible to detect the cases of frequency difference by examining the orbits of the embedded signals, as there would be a greater number of orbits in a higher frequency attractor, but clearly caution must be used in the interpretation of the asymmetries.

The situation becomes more complex with the addition of the time lags between signals, and the assumption that these lag times indicate a directional relationship in the direction from lead to lag (Meeren et al., 2002). I am hesitant to accept the reasoning as there exist examples that can show lead-lag relationships of this type, yet show the strength of prediction being higher in the opposite direction. Indeed, it is possible for both signals to have their $Dv_{\max}$ to be leading the other or both lagging due to the

asymmetry of the divination function, and thus only when the $Dv_{\max}$ is significantly above chance levels and the leads are symmetric and opposite would I feel justified in claiming that a leading-lagging relationship exists.

Since the periodic component is unaffected by the lead-lag relationships any lead discovered by the divination function is not prone to an error that can be made with other measures – with two periodic signals a maximum cross-correlation could be interpreted as a short lead in one direction or a long lag in the other; the periodic component of the divination function would be unaffected by such issues of phase, so we are safer in accepting the directionality based on a lead or lag than we would be if we were using a maximum of the cross-correlation to identify the lag, for example.

The last point about the divination function itself is that it is intuitive – the embedded points, with their multiple coordinates, represent snapshots of the dynamics over a short period of time, with the 2 dimensional case being roughly interpretable as a slope, and the higher dimensional cases describing more complex behaviours or patterns. A high Dv is interpreted as being a strong association of specific patterns on one attractor with specific patterns on another, essentially being the average of the ability to predict from one attractor to the other based on these patterns. With a number of neighbours selected it becomes a matter of knowing where roughly on an attractor one signal is from knowing roughly where the other is, essentially a measure of the synchrony of states, no matter how different those states are. The divination coefficient (or function) thus has a simple interpretation, that it quantifies the ability to predict the state of one system from the state of a second system.

For my divination program the selection of the sampling and delay were automated to create a systematic approach to their selection, yet the values of all the parameters used in the analysis of the data were similar to those used by Arnhold et al. (1999), who investigated EEG data from epileptic patients, suggesting that the nature of the data constrains these values. In particular, the embedding dimension, delay, number of embedded points per second and the length of the sample represented by a vector are all similar between the two methods.

Since the parameters were arrived at through different methods they are likely good estimates of the appropriate values, and the systematic approach has the advantage that it allows for automation of the computation. Automation also allows the selection of a delay and sampling for each signal, allowing some optimization of the embedding, rather than determining a single set of values and applying these to all the data.

Arnhold et al. (1999) used a dimension of 10 for their method while the false nearest neighbour method suggested that 6 dimensions were appropriate for my embedding. Some differences exist between the two paradigms, the main difference being that Arnold et al. were working on human EEG, while I am using recordings from depth electrodes in rats. The species may have differences in the dimension of brain activity, and this is the value for which the two methods differ most significantly. As well, the FNNP method may underestimate the true dimension, as we are taking the dimension for which the FNNP fell below 1%; this under representation is not problematic, if fewer than 1% of the points are misrepresented it will have minimal impact on the $Dv_{\max}$.

The delay chosen by the autocorrelation decay for the seizure data fell in the range of 29-101 ms, but was characteristically near 40-60 ms, for spontaneous recordings the

range was 23-41 ms, with an average of 31.45 ms; Arnhold et al. (1999) used a delay of 5 time units, which corresponds to $5/173$ of a second, or ~ 29 ms. We see that this is very close to the value for the non-seizure recordings from the rats; since Arnhold et al. determined that 10 dimensions were the appropriate embedding dimension we can assume that they were not measuring this on seizure data, which is known to be as low as 2-4 dimensions, and so we see that the value is about equal.

The number of embedded points per second determines how well represented the attractor is, and should be sparse enough to allow for finding neighbours on other orbits, yet dense enough to properly represent the attractor. Arnhold et al. (1999) sampled at 173 Hz and used a decimation factor of 5 to reduce the amount of data they were using, this results in 34.6 embedded points per second; my samples were recorded at 1000 Hz, and the sampling was between 7 and 23, with an average sampling of about 16 resulting in about 60 embedded points per second.

Lastly, the length of signal represented by the embedded points is $(m-1)\tau$, so the embedded points represent 261 ms in Arnhold et al. (1999). Because I have selected a delay based on the signal's autocorrelation function this length varies for the case studied, with a length of 150 ms or so for the non-seizure data and a length of roughly 250ms for the seizure data.

The similarity of the values calculated by my program to those determined by Arnhold et al. (1999) is encouraging, as it suggests that the methods I have used to evaluate appropriate parameters are close to those determined to be “optimal” for a similar set of data.

Significance of the divination function

One issue mentioned earlier in our results is that it becomes important in some instances to judge the significance of a given divination value. What constitutes a “high” value, and how can we compare between values generated by different attractors? A few attempts were made to judge the significance of the divination values, all of which depend essentially on the creation of surrogate data. The nature of the surrogate data determines what null hypothesis is tested, and as we have no knowledge of the distribution we use Tchebychev’s inequality to determine the probability of a given result occurring by chance. Tchebychev’s inequality is a very conservative statement that for any distribution, at most $1/k^2$ of the points are found k or more standard deviations from the mean. We can therefore achieve a significance of .05 by rejecting the null hypothesis that the sample data is from the same population as the surrogate by rejecting the null hypothesis when the value exceeds the mean by 4.472 standard deviations.

The first surrogates constructed were simply time-shifted versions of the reference signal, shifted by a delay greater than 400 ms but less than 800 ms. In the shift-surrogates the same patterns of activity are seen (as the recorded signals are relatively stationary over these delays), and the frequency distribution should be similar, theoretically only breaking the temporal relationship between the signals. Any difference could hopefully be attributed to the decoupling of the signals temporally. One difficulty that was unforeseen was the effect of insensitivity to the phase of periodic signals on the shifted surrogate signals. As the seizure data has a strong quasi-periodic component, the decrease in the $Dv_{\max}$ when a time shift is applied is much less than it would be for a non-

periodic signal, and this is thus a poor way to detect whether the value is significant, as we have not severed the connection of the periodic portion of the signals.

A second attempt to construct surrogates was the randomization of the embedded points on the reference signal. This would result in a random referred neighbour being chosen for each point of the sample signal, and the null hypothesis would thus be that the states of the two systems are related only by chance, with the alternate hypothesis that the relationship is more than chance would suggest. This seems a very weak statement, but it is nonetheless a method that addresses the difficulty with periodic signals. It was apparent after some consideration that this was quite similar to simply choosing to associate random embedded points from the sample system with the nearest-neighbour points, resulting in a measure of the average ratio of the nearest-neighbour distance and the average distance between random points.

The method involving randomization of the embedded points must be done for each sample attractor, as the dynamics of the attractor determine the relationship between the neighbour distance and the average distance, and a large enough sample would need to be taken in each comparison to generate a standard deviation for the Tchebychev inequality.

A third attempt to construct surrogates is based on the results of predicting noise from noise, and using the values of noise-noise divinations as the surrogate values. The embedded points of noise could be used with the second method as a way to simply randomly select neighbours for the points within the embedded signal, but as was mentioned, this involves recalculating the values for every embedded signal tested. Since the divination coefficient is a measure of the ratio of a nearest neighbour distance to a referred neighbour distance let us examine how these values change. In both cases the

referred neighbour distance is a distance between points selected roughly at random on the attractor, thus it is a measure of the range of the attractor. The average distance between random pairs of points in noise occupying a hypercube that bounds the signal should be similar in magnitude to the average distance between random points of the signal. Since the attractor is presumably less diffuse than the noise (as it is embedded in a suitable dimension), the nearest neighbour distances on the signal are likely to be much smaller than the nearest neighbour distances on noise. We thus expect the values obtained for a noise-noise divination to be larger than those obtained for the signal-noise divination. The value obtained from noise-noise comparisons is thus more conservative than those obtained from the use of the permuted signal, and thus we can adopt this as a method of judging significant synchronies.

The noise-noise values do not allow for easy comparison between the values obtained on different signals, as the attractor dynamics influence how high above noise a significant divination value will be. To compare the divination values we must eliminate the effects of the attractor dynamics on the value – this can be done, and I will show that an ordinal divination value solves this problem. For the current situation we can use average and standard deviations calculated from samples of noise to determine if the signals are more significantly synchronised than chance would suggest, and a horizontal line has been drawn in Figure 19 as well as in the plots of $Dv_{\max}$ in appendix A indicating 4.472 standard deviations above the $Dv_{\max}$ for two signals of random noise.

Dimension in biological data

The dimension program was written in order to determine appropriate embedding dimensions for the biological data, and as a companion program for the divination function program, as we need to have a dimension selected in order to perform the embedding. As the results of the tests on simple systems and noise indicate, it accurately determined the appropriate embedding dimensions for a variety of known signals. The surprising result was that when applied to the data from the pre- and post-kindled rats it found different embedding dimensions, suggesting a change in the dimensionality of the brain activity.

Figure 9 shows the false nearest neighbour percentage for increasing embedding dimension for the data from the rats in the immobile condition. The difference in the slopes of the line representing the drop in FNNP with dimension for the pre-kindled and post-kindled rats is immediately apparent, and as would be expected by such a noticeable difference, it is statistically significant, with a $t = 5.06$ and $p < .001$. The results in Figure 10 are equally obvious, showing a clear difference between the slopes of the FNNP vs. dimension lines for the pre- and post-kindling conditions in the sniffing rats, and it is also statistically significant, with $t = 3.82$ and $p < .001$. There was insufficient evidence to suggest a difference of dimension between the Slow and the Fast rats in either case.

Dimension has been used as a method to identify epileptic seizure and to attempt to predict seizure. It is known that during epileptic seizure the human brain drops from as high as 10 dimensions to a mere 3-4 (King, 1991). Dimension based methods have been used in seizure prediction with some success, detecting pre-ictal changes in dimension as early as 10 minutes before seizure onset (Lehnertz et al., 2003). The fact that kindling

produces a loss of dimension in rats is interesting as a new finding, and has a few possible explanations.

One possibility is that the brains of the kindled rats have become simply more variable, with random firings of neurons and a reduction in the apparent dimensionality, similar to low dimensional chaos. Some studies have suggested that during epileptic seizures brain activity is a result of low dimensional chaos (Pijn et al., 1991), and it could be that the rats are essentially experiencing chaotic firings due to the creation of new attractors.

Another possibility is that just as we see decreases in dimension pre-ictally in humans (Lehnertz et al., 2003) that a similar sort of pre-seizure state exists in kindled animals, signifying a greater tendency to slip into the basin of attraction of seizure activity. The pre-ictal drop in dimension found before seizure may not be a guarantee of seizure, but merely a state that is closer to seizure, and renders the epileptic prone to seizure, thus “predicting” it; if so there is a strong parallel to the kindled rat, as they have a lowered discharge threshold and a tendency to suffer spontaneous seizures, and kindling may just be a permanent resetting of the brain’s dynamics to a pre-seizure state.

The dimension of the rat brain during stimulated seizure was even lower than that of the kindled rats, suggesting that it is similar to the state in human seizure, a drop to very low dimension due either to noise or a chaotic attractor. The lowest averages were always found in the stimulated area, with the stimulated amygdala scoring the lowest of all areas, at 3.91 dimensions. This seems high compared to the values reported in literature of 2-4 dimensions during seizure (Pijn et al., 1991), but it should be noted that this was the first set of seizures elicited in un-kindled rats, and that the lower figure of 2-4 dimensions was obtained from kindled rats. It is possible that the seizures are of lower dimension once

fully kindled due to the propagation of the faulty networks that result in the chaotic neuronal firings. Another possibility stems from the method of measure; the most common measure of dimension is the correlation dimension, while we are using a geometric method to find the minimum embedding dimension. This method may tend to overestimate the dimension of low dimensional chaos as we know it is increased by the presence of “noise”, and it may underestimate the dimension of the higher dimensional systems as it is finding the minimum embedding dimension, and thus will treat small numbers of errors as negligible provided the majority of points are not false nearest neighbours.

Another possibility is unfortunately that the drop in dimension is not due to the kindling but to some other effect, such as aging. Since the focus of my thesis was the development of a method to measure synchrony the testing was preformed on data gathered in the kindling of rats for another experiment, and the data did not include control animals. While it is unlikely that aging in rats causes a large and consistent loss of dimension, it is not impossible, though one suspects given the magnitude of the effect that it would have been commented upon in studies that did use controls.

The fact that no difference was found between the Fast and Slow rats in either pre- or post-kindled groups is not worrisome, as there are many explanations. It is possible that there are no differences, and that the susceptibility to kindling is due to a separate mechanism, or that the difference is too small to detect by this method, or even that the sample sizes are too small to detect the difference.

Divination in biological data

Spontaneous recorded activity

The most important of the results of the testing of the divination function on biological data was the finding that pre-kindling differences exist between the Slow and Fast rats, and that these differences disappear with kindling. The differences occur in both the immobile and the sniffing condition, and were statistically significant with $p < .05$.

The disappearance of the difference in both cases is a nice result, as it suggests that kindling is driving both Fast and Slow rats to the same endpoint, a brain with a reduced threshold for seizure, and that the neurodynamic changes that accompany kindling are common to both strains. Of course, given the small sample size used in the testing (6 Slow and 5 Fast rats) it is possible that a difference still exists, and that we simply have too small a sample to reveal it statistically.

The pre-kindling difference depended on the state of the animal, the Slow rats having a lower $Dv_{\max}$ than the Fast rats in the immobile state, but a higher $Dv_{\max}$ in the sniffing state. This suggests that the issue is not something as simple as a less or more synchronous system, but that differences in the ability to synchronize groups of neurons may be a factor. Epileptic seizure is characterised by large amplitude spike/wave rhythms in the EEG, and such firings indicate synchronies between large numbers of neurons in the region of the electrode. If building synchronies that grow out of control are responsible for the mass firings of neurons in epileptic seizure a greater ability to control the synchronicity between groups of neurons and brain regions may allow the early diffusion of potential seizures, explaining the resistance of the Slow rats to seizure.

The post-kindling levels of $Dv_{\max}$ are very similar (.309043 and .306604 for sniffing and immobile respectively) in both the sniffing and immobile recording conditions, in addition to not being different between the Slow and Fast rats in each condition. The differences in $Dv_{\max}$ that normally occur between the immobile and sniffing recordings have been reduced in both conditions, implying perhaps a more restricted range of $Dv_{\max}$ being attainable in the kindled rats. While it is purely speculative to hypothesize that a new similar state has been induced in both strains of rat, the evidence suggests that the strains did differ in inter-region predictability prior to kindling, and that this difference is not detectable after kindling. This homogenizing of the inter-region synchronies may be due to the induction of similar dynamics in all animals, such as a pre-seizure state in which the animals are more prone to seizure and cannot establish the range of states required to escape the basin of attraction of the seizure or pre-seizure states, and a state that was in some ways similar to a pre-seizure state could explain both the drop in dimension observed and the change in $Dv_{\max}$.

Stimulated after discharges

The stimulated seizures were examined from a qualitative perspective, in order to test whether the $Dv_{\max}$ was capable of detecting changes in the generalised synchrony that may occur within the course of a seizure. The results indicate that the divination function program is capable of detecting such changes, and as well has shown the detection of shifts in the generalised synchrony between regions following seizure offset, some of which may be predicted by changes in the $Dv_{\max}$ lag.

As was noted earlier, the number of changes that occur in a single seizure are too many for me to address in this thesis, but the generalities of the seizures are worth mentioning.

As would be expected, a high value of $Dv_{\max}$ occurs when two regions are in activity that is similar in appearance – this type of activity “locking” is seen between the left and right hippocampus during the seizure, and between the ipsilateral amygdala-hippocampus pairs toward the end of the recording, and is clearly visible in Figure 20 as periods in which activity on the various electrodes is quite similar, as between the electrodes from the left hippocampus and left amygdala after the 25000th data point. The $Dv_{\max}$ for the post-seizure period is still abnormally high when compared with the resting values, and synchronies are still clearly disrupted, with very strong correlations between the regions. Low values of $Dv_{\max}$ are seen when two regions are clearly not following similar activities, such as between the amygdalae and the hippocampi during the large amplitude seizure at the beginning of the recording. We would expect these values to be low, as the hippocampal activity seems to be quite distinct from the activity in the amygdalae, and is of a much different frequency and form.

In many of the seizures a drop in $Dv_{\max}$ is observed during the seizure between areas, particularly areas not involved in the seizure. The disruption caused by the seizing brain regions may sever communication between other regions, and thus reduce the degree to which regions are predictive of one another.

That a change in the $Dv_{\max}$ lag can precede the offset of the large amplitude spiking of the seizure is an interesting possibility, as it may suggest that the driving mechanism is becoming derailed between the two regions. Given that the seizure was stimulated no investigation can be made of how the seizure began, but the offset shows that the two regions that were most in synchrony with each other have the lag between them disappearing prior to the seizure offset. If we do choose to interpret the lag as indicating

that one region is driving the other, then the left hippocampus is in this case driving the right, and it would seem then that the seizure is being driven from the non-stimulated hippocampus. Whether the loss of the lag between the regions is a symptom or a cause of the seizure offset is not evident.

Here the difficulty of using a window to analyse the data becomes apparent, for every result is an average from a window 3 seconds long, and the state of the brain is changing quite rapidly during seizure offset. We have readings of the $Dv_{\max}$ and lag every 200 ms, but while this can help to identify when an event takes place it comes as no surprise that the width of the smaller increases in the $Dv_{\max}$ occur over a minimum of about 15 windows, as 15×200 ms is our 3000 ms window length. This result is analogous to sliding a toboggan along over a path with one lump; the average height of a passenger on the toboggan will increase during an arc of roughly the length of the toboggan, with its peak at the site of the bump – likewise, we must attempt to identify the timing of events within these windows by this “toboggan” method, and the resolution is finite. We cannot use windows below a certain length, as we need several orbits of the attractor in order to find meaningful neighbours, and with the lower frequencies that we are interested in lying in the range of 2-4 Hz we must have at least $(j+1) \times 500$ ms of data if we wish to use j neighbours in our comparison in order to have neighbours selected that are not temporally proximate. The lags are likewise an average for the area of the window, and thus if they increase or decrease smoothly they are likely representative of the entire window, but when they fluctuate rapidly from window to window it can be difficult to interpret over what periods the signals lag and lead one another.

The results of the testing of the divination function on the seizure data are intriguing in that they indicate that much synchrony exists in the post seizure period and that there continue to be changes in the communication between these regions following the seizure offset. This is encouraging as the goal of the program is to detect minute changes in synchrony between signals, and it has proven its ability to do this and thus to be of use in neuroscience research.

Changes to the divination function

As it stands the divination function program is capable of detecting the presence of synchronous events between signals, and of doing so with simulated delays between the signals in order to find the lag between signals. The value of the divination function can be thought of as a measure of the strength of the predictability from one attractor to another, in essence a measure of the ability to map one section of the phase space to a section in another phase space. In this way it is similar to the concept of mutual information, though non-symmetric and operating on the embedded vectors. The possibility of simply using a mutual information approach on the embedded vectors seems at first glance to be a simple way to improve on the method, but it is unfortunately not feasible. To use mutual information we must sort the values into “bins”, which is not an easy thing to accomplish on an attractor, as the points exist spread through space. The concept of a bin is like the concept of a neighbourhood, but while our neighbourhoods may be about any point in the embedding, the bins must remain fixed, and to classify a point as falling into a bin is often to ignore its greater proximity (and thus more similar state of the system) to neighbouring points.

Given that we cannot replace the neighbour method without difficulty, we have to determine what changes can be made to improve the method and the program. The method of “sampling” the embedded time series to find neighbours that are not related temporally is the first change I would make to the program. While sampling does achieve the result that the points selected as neighbours are unlikely to be close in time to the sample point it results in less than ideal neighbours. If for example the best neighbour falls in between the sampled points we are making a less accurate prediction. An improvement to the method would be to determine the minimum time (s) that must separate the sample point and a neighbour, as we have described in the methods section. Once this is done, we search for the nearest neighbour among points more removed than s , and having identified it, we then remove from consideration all points within s of the first neighbour before searching for a second neighbour. This ensures that each neighbour selected will be the optimal neighbour from an orbit.

There are a few problems with this method, one of which is the presence of orbits that occur once. A transient event such as an isolated spike may have no useful neighbours, and if chosen as a point for the calculation of s the value may be very large, as it must first eliminate every point on the orbit. It would thus make sense to determine a percentage of points which must pass the test of having their nearest neighbours s units away, and this may prevent inflation of the value of s . Given that s is properly estimated the majority of points will now be selecting as neighbours points that are on other orbits,. We must be certain not to choose to use more neighbours than we have regular orbits, or we will be choosing neighbours that are not distant in time, and thus are related through temporal proximity rather than similar dynamics. Indeed, with s as the minimum number

of data points separating the sample point and its neighbour we could select for each point as many neighbours as we can without selecting a point distant from the sample point or any neighbour by $s+1$ points. In this way, rather than choosing a number of neighbours to use in the calculation all neighbours will be used that fall on points that are neighbours and on different orbits.

When the reference point is chosen it may have fewer neighbours than the sample point had – if this is the case, we use the minimum number of neighbours, so that if our sample point has j neighbours and our reference point k neighbours ($k < j$), we will simply use the first k of our reference point's neighbours. This will increase the computation time for the neighbours a great deal, as it requires s times as many comparisons per neighbour calculated, but would result in much better matches, and prevents the possible problem of selecting more neighbours than there are temporally removed neighbours.

Doing this will increase the number of points available for the calculation of the divination function, but a fraction of the points can be randomly selected from the embedded points, as provided they are sampled randomly they should be representative, much as we were able to use $1/s$ of the points to get an accurate value.

An ordinal divination function?

An idea for a modification of the divination function occurred to me during the last phases of testing the program that I had written to compute the divination function.

Rather than base the strength of the prediction on the distances between the sample point and the referred and nearest neighbour, it is possible to use ordinal version, based on the neighbour relation between the points. Such a measure could still detect whether a given

referred neighbour was a close match, but rather than be affected by how proximate the nearest neighbour is, the calculation of divination coefficient would be a comparison of their neighbour status. If the k^{th} neighbour is given the value k , and we subtract the referred neighbour value from the nearest neighbour value we get a number between 0 and $N-1$, where N is the number of points in our sample. A perfect match would result in the referred neighbour being the nearest (1^{st}) neighbour, and as such would be a value of zero. This simplifies the calculation in many ways, and provides a clearer portrait of the expected behaviour of the divination coefficient.

The expected value of randomly associated (non-coupled) systems would no longer depend on the precise distances between points and thus the distribution of the time series values; using the original cardinal divination coefficient we had to employ fairly sophisticated methods to get an expected value for uniform noise, and it did not provide us with an easily calculated standard deviation or error. The calculation would have been far more complex for even a Gaussian distribution, and to have a distribution of values similar to the sample being tested required empirical methods, such as the construction of shift surrogates or random point permutations to assess whether the divination function was exceeding the expected level for randomly associated signals. This severely limited the ability to judge the significance of a given value of divination coefficient, and led to our use of comparisons between values with ANOVA.

An ordinal approach removes any necessity to construct surrogates or worry about the distribution of values in the time series – all time series are essentially reduced to a table of neighbour relationships, and we can calculate the expected value of a randomly associated set of points by simply knowing the number of points involved in the

comparison. With N points in each signal, randomly associated, the ordinal divination coefficient with 1 neighbour would on average find that the nearest neighbour was associated with the $((N-1)/2)^{\text{th}}$ neighbour, and we can then say that the average value would be $(3-N)/2$. Since the worst case would be to associate every nearest neighbour with the $(N-1)^{\text{th}}$ neighbour, the minimum value thus becomes $2-N$, and we can scale this so that all values lie between 1 and -1 by adding $(2-N)/2$, and dividing by $(2-N)/2$. The average value for noise (uniform, Gaussian or otherwise) would then be 0, with a perfect prediction resulting in a value of 1.

Some other problems are eliminated with the use of the ordinal divination coefficient – we ran into some division by zero errors when for example there were multiple copies of a point in an embedding, as the nearest neighbour distance (and the referred neighbour distance) can take on a value of 0 – since there is no longer a division performed, the existence of multiple points no longer poses a problem for the ordinal calculation. The issues associated with particularly good nearest neighbour matches resulting in very small divination coefficients is also solved – while the cardinal divination coefficient is susceptible to this kind of difficulty due to the division of the very small nearest neighbour distance by a much larger referred neighbour distance, the ordinal divination coefficient is sensitive only to how closely ranked the referred neighbour is. This solves the odd (and counter intuitive) result that predicting a sine wave from noise produces a lower cardinal divination coefficient than does the prediction between two time series of noise – since the dynamics of the embedded sine wave no longer enter into the equation, the divination coefficient simply becomes the random selection of a match to the first

neighbour, resulting in an expected value of 0, the same as for a noise-noise prediction. This seems correct from an intuitive point as well, that any signal predicted from noise should have the same expected divination.

The cardinal divination coefficient is also sensitive to certain other changes – even with $Dv(A,B) = 1$ and $Dv(B,A) = 1$, indicating a perfect prediction in both directions, $Dv(A,C)$ and $Dv(B,C)$ can have different values, due to the differing neighbour distances in A and B, though $Dv(C,A)$ must in this case be equal to $Dv(C,B)$, as the distances are all measured on the same attractor, and both A and B share a neighborhood matrix. This strange situation cannot occur with the ordinal version of the divination coefficient, as the neighbourhood structure of A and B are identical (at least for the number of neighbours involved and at that dimension) and thus $Dv(A,C)$ and $Dv(B,C)$ would be equal.

The idea of the ordinal divination coefficient can be extended in other ways – logically, the divination coefficient is simply the average of a set of neighbour differences based on the comparison – by extension, other values that can be calculated from this would be the other moments about the mean, the variance, skewness and kurtosis of the neighbour differences. It is not initially obvious what the interpretation of all of these would be, but we know for example that although the prediction of noise from noise would have the same average (and thus ordinal divination coefficient) as the prediction of a sine wave to a sine wave of one half the frequency, the distribution of these values would be very different, with the noise-noise distribution being uniform (equally likelihood of all differences from 0 to $2-N$) and the sine-sine distribution being strongly bimodal, with half the values being near 0, and half being near $2-N$. Can we say that being able to predict well half the time and abysmally half the time is any better than having no prediction at

all? It would be hard to rate one of these as a “better” prediction, yet they clearly are different, and the other moments of these neighbourhood relations can clarify these differences.

While there are many advantages to the use of an ordinal rather than a cardinal approach to the neighbour comparisons, there are some issues, and these should be considered before deciding to use one or the other. Firstly, there are longer computation times associated with the use of the ordinal method – rather than having to only find a first or first k neighbours for each point we must find the full set of neighbour relations for all points – this involves increasing the number of comparisons from the $N-1$ or $k(N-k/2-1)$ per point required for the cardinal method to the $N^2/2$ comparisons needed per point for the ordinal method. When the number of points is large this can be thousands or millions of times the number of operations, although with very large N there are faster algorithms than the brute force methods I used in this program.

Secondly there is an issue of interpretation – the cardinal method is very influenced by changes in the nearest neighbour distance, especially as it relates to the referred neighbour distance; if these variances are due to the nature of the attractor it may make sense to use the cardinal method, while if they are due to sampling over the attractor it is likely best to ignore the actual distances in favour of the rank order.

A third consideration is the type of signals/attractors we are dealing with; in the case of the biological data examined here, we often have signals with both small high frequency components and lower frequency large amplitude components; the different distances in the parts of our attractor may cloud the qualitative predictive power between such signals. On the other hand, the sensitivity of the cardinal method to distances prevents us from

erroneously reporting high values of divination spuriously, such as when the number of neighbours chosen for the analysis exceeds the number of orbits present in the sample – a referred neighbour being the 5th nearest neighbour is meaningless if there is only one “true” neighbour, and thus the cardinal method prevents misattributing these coincidences as meaningful predictions.

Overall, the ordinal divination coefficient seems like a wise choice, provided it is used with the neighbour searching algorithm describe in the *Changes to the divination function* section. As the method still relies on using the actual distances in some phase space it is not a purely ordinal method; this prevents us from using neighbours that are unrelated. The finding of neighbours and their ordering is all performed in an embedding, with some metric, thus providing accurate evaluation of which neighbours are closest, and evaluates only as many neighbours as are on nearby orbits. These two changes taken together will allow much simpler comparisons between signals, and will also allow the evaluation of precise probability tables that detail when a given value of the divination coefficient is significant, based not on empirical testing of data, but on theory. The ability to state in precise terms what the probabilities are of getting any level of divination coefficient will improve the ability to detect significant increases in the divination coefficient as the conservative empirical system will not be needed.

Future directions

One of the long-term goals of the development of this method is to apply it to the prediction of epileptic seizure. The divination function has properties that set it apart from many of the techniques that have thus far been employed, and makes use of the

properties that have thus far been the most successful in seizure prediction. Seizure prediction at this time is done on data recorded in the lab, and subsequently analysed, due to the lengthy computation times of some of these methods. One goal shared by most researchers is increase the speed of the calculations to the point that they could be applied in real time, and serve as a warning of seizure onset, or even as a trigger for methods of seizure prevention.

One obvious future direction would be to attempt to increase the speed of the divination calculations to real time. This is unlikely to be achieved simply by application of different algorithms and increases in computer speed, at least at present, as the reduction in computation time required to achieve this goal is on the order of 60-100 times. It is thus important to identify how one can reduce the amount of data to a more manageable quantity, while maintaining those elements required for prediction. Near optimal techniques have been developed for the rapid calculation of neighbour relationships, but are of little use in the relatively small data sets that would be used in seizure prediction. I suggest thus that it is not the application of more intricate algorithms that is needed, but ways in which the data can be reduced, thus reducing the amount of data that remains to be analysed.

As multiple methods can be combined, the use of fast methods such as simple linear correlations, mutual information and spectrum analysis may together with stripped down versions of the divination coefficient provide a faster and still insightful analysis.

One method to reduce the amount of data would be to involve some form of cluster analysis; rather than treat each state of the system as a point in phase space we would reduce the attractor to a finite number of clusters of proximate points, and treat these

points as a whole, or simply count each point as being part of a given neighbourhood.

Cluster analysis is a large area of research in mathematics, and while none of the methods I examined have seemed appropriate, the concept of clustering the states to reduce the number of comparisons seems like an appropriate way to conserve the relationships between states while limiting the computation time.

One possible clustering technique would be to base the clustering off the algorithm described in *Changes to the divination function*, and to build each cluster by finding a neighbouring point that falls beyond a certain temporal range – all such points would constitute one element of the cluster, along with the points within the range previously used. Given the values of τ at which the autocorrelation function drops below $1/e$ for the cases examined we could reduce the number of points by a factor of 20-30, attaining computation times near real time, though testing would need to be done to ensure that the method still provides meaningful results on the limited number of points that remained.

The use of the program to determine embedding dimension on the data from the rats before and after kindling uncovered a gap in the literature of epilepsy, in that control groups are nearly never used in the investigation of dimension in epilepsy. The reduction of dimension during epileptic seizure is well documented, as is the smaller drop in dimension accompanying the pre-seizure period, but comparison of the dimension of brain activity in controls is lacking. It would be interesting to know whether epileptic patients also show a lower dimension than do controls, and if not, whether specific types of epilepsy have differing dimensions for non-ictal activity. Since the kindled rats show a decrease in dimension of activity it is important to know whether epileptic patients show the same lower level of activity compared to controls in order to interpret the

change, and it seems that this has not been investigated. Investigation of the non-ictal state in epileptics as a comparison to controls would therefore be a valuable contribution to the field.

Finally, we can obviously use the programs developed to further investigate both stimulated seizure and changes due to kindling; these topics were addressed briefly as part of a test of the programs on biological signals, but much more could be done on either of these subjects.

The seizures were examined only briefly, due to the immense amount of information provided by the recordings and the variety present in the seizures. Although we have 22 stimulated seizures, the recordings vary in length, duration of the seizure, the brain regions involved in the seizure and so on, in such a way that makes any averaging impossible, nor can comparisons be made between seizures as the different samplings (s) used make this largely impossible. Classification of the changes visible in the seizure recordings may allow a statistical approach, by correlating changes in the electrode recording with events on both the record of lags and the $Dv_{\max}$ for the seizures. This was done to some extent for the one seizure examined in the results section, but a larger analysis may reveal more features that are shared among the seizures, and perhaps lead to a better understanding of the events leading up to seizure offset, or in the post-ictal period.

The changes in $Dv_{\max}$ observed in the rats that underwent kindling were surprising, and given that a significant difference was found with a small sample doubtless more remains to be discovered. Future work could also examine the kindling process to determine at what point the dimension begins to change for the different strains (Fast and Slow), and

as well could include a control group, to verify that the changes are not a result of aging or other factors. In addition, there were trends indicating that the changes in $Dv_{\max}$ were not equal in all areas, and with a larger sample it might be possible to determine in what areas the changes were occurring.

To conclude, I think that the divination coefficient has proven itself effective in the detection of the forms of synchrony it was designed to detect, and that the existing method could be further used in research into the nature of seizure and the changes involved in the kindling model of epilepsy. Some small changes to the method and program could enable much easier comparisons between $Dv_{\max}$ on different signal pairs, and the added precision in determining significance will greatly facilitate the detection of significant generalized synchronies.

References

Andrzejak, R.G., Mormann, F., Kreuz, T., Rieke, C., & Kraskov, A. (2003). Testing the null hypothesis of the nonexistence of a preseizure state. *Physical Review E*, 67, 10901.

Arnhold, J., Grassberger, P., Lehnertz, K., & Elger, C.E. (1999). A robust method for detecting interdependences: application to intracranially recorded EEG. *Physica D*, 134, 419-430.

Bhattacharyya P., & Chakrabarti B.K. (2003). The mean distance to the n-th neighbour in a uniform distribution of random points: An application of probability theory. Retrieved November 23, 2003, arXiv:math.PR/0212230 v2.

Casdagli, M., Eubank, S., Farmer, J.D., & Gibson, J. (1991). State space reconstruction in the presence of noise. *Physica D*, 51, 52-98.

Casdagli, M.C., Iasemidis, L.D., Savit, S., Gilmore, R.L., Roper, S.N., & Sackellares, C. (1997). Non-linearity in invasive EEG recordings from patients with temporal lobe epilepsy. *Electroencephalography and Clinical Neurophysiology*, 102, 98-105.

Dayan, P. & Abbot, L.F. (2001). *Theoretical Neuroscience*. Cambridge, MA:MIT Press.

Dudek, F.E., Patrylo, P.R., & Wuarin, J.P. (1999). Mechanisms of neuronal synchronization during epileptiform activity. In: A.V. Delgado-Escueta, B. Wilson, R. Olsen & R.J. Porter (Eds.), *Jasper's Basic Mechanism of the Epilepsies 3rd Ed., Advances in Neurology Vol. 79*, Philadelphia: Lippincott-Raven.

Glass, L. (2001). Synchronization and rhythmic processes in physiology. *Nature*, 410, 277-284.

Goddard, G.V. (1967). Development of epileptic seizures through brain stimulation at low intensity. *Nature*, 214, 1020-1021.

Hegger, R., Kantz, H., & Schreiber, T. (1999). Practical implementation of nonlinear time series methods: The TISEAN package. *Chaos*, 9, 413-435.

Iasemidis, L.D., Sackellares, J.C., Zaveri, H.P., & Williams, W.J. (1990). Phase space topography and the Lyapunov exponent of electrocorticograms in partial seizures. *Brain Topography*, 2, 187-201.

Jerger, K.K., Netoff, T.I., Francis, J.T., Sauer, T., Pecora, L., Weinstein, S.L., & Schiff, S.J. (2001). Early seizure detection. *Journal of Clinical Neurophysiology*, 18, 259-268.

Kantz, H., & Schreiber, T. (1999). *Nonlinear Time Series Analysis*. Cambridge: Cambridge University Press.

Kennel, M.B., Brown, R., & Abarbanel, H.D.I. (1992). Determining embedding dimension for state space reconstruction using a geometric construction. *Physical Review A*, 45, 3403-3411.

King, C.G. (1991). Fractal and chaotic dynamics in nervous systems. *Progress in neurobiology*, 36, 279-308.

Lange, H.H., Lieb, J.P., Engel, J., Jr., & Crandall PH (1983). Temporo-spatial patterns of pre-ictal spike activity in human temporal lobe epilepsy. *Electroencephalography in Clinical Neurophysiology*, 56, 543-55.

Lehnertz, K. (1999). Non-linear time series analysis of intracranial EEG recordings in patients with epilepsy – an overview. *International Journal of Psychophysiology*, 34, 45-52.

Lehnertz, K., Andrzejak, R.G., Arnhold, J., Kreuz, T., Mormann, F., Rieke, C., Widman, G., & Elger, C.E. (2001). Nonlinear EEG analysis in epilepsy: Its possible use for interictal focus localization, seizure anticipation, and prevention. *Journal of Clinical Neurophysiology*, 18, 209-222.

Lehnertz, K., & Elger, C.E. (1995). Spatio-temporal dynamics of the primary epileptogenic area in the temporal lobe epilepsy characterized by neuronal complexity loss. *Electroencephalography in Clinical Neurophysiology*, 95, 108-117.

Lehnertz, K. & Elger, C.E. (1998). Can epileptic seizures be predicted? Evidence from nonlinear time series analyses of brain electrical activity. *Physical Review Letters*, 80, 5019-5023.

Lehnertz, K., Mormann, F., Kreuz, T., Andrzejak, R.G., Rieke, C., David, P., & Elger, C.E. (2003). Seizure prediction by nonlinear EEG analysis. *IEEE Engineering in Medicine and Biology*, 22, 57-63.

Lehnertz, K., Widman, G., Andrzejak, R., Arnhold, J., & Elger, C.E. (1999). Is it possible to anticipate seizure onset by non-linear analysis of intracerebral EEG in human partial epilepsies? *Revue Neurologique*, 155, 454-456.

Lerner, D.E. (1996). Monitoring changing dynamics with correlation integrals: Case study of an epileptic seizure. *Physica D*, 97, 563-576.

Le Van Quyen, M., Adam, C., Baulac, M., Martinerie, J., & Varela, F.J. (1998). Nonlinear interdependencies of EEG signals in human intracranially recorded temporal lobe seizures. *Brain Research*, 792, 24-40.

Le Van Quyen, M., & Martinerie, J. (1999). Characterization of spatio-temporal changes prior to epileptic seizure by a measure of dynamical similarity. *Theory in Biosciences*, 118, 199-208.

Le Van Quyen, M., Martinerie, J., Navarro, V., Baulac, M., & Varela, F.J. (2001). Characterizing neurodynamic changes before seizures. *Journal of Clinical Neurophysiology*, 18, 191-208.

Le Van Quyen, M., Martinerie, J., Navarro, V., Boon, P., D'Havé, M., Adam, C., Renault, B., Varela, F., & Baulac, M. (2001b). Anticipation of epileptic seizures from standard EEG recordings. *Lancet*, 357, 183-188.

Lewkowicz, M., Levitan, J., Puzanov, N., Shnerb, N., & Saermark, K. (2002). Description of complex time series by multipoles. *Physica A*, 311, 260-274.

Malmivuo, J. (1999). Theoretical limits of the EEG method are not yet reached. *International Journal of Bioelectromagnetism*, 1(1), 2-3.

Martinerie, J., Adam, C., Le Van Quyen, M., Baulac, M., Clemanceau, S., Renault, B., & Varela, F.J. (1998). Epileptic seizures can be anticipated by non-linear analysis. *Nature Medicine*, 4, 1173-1176.

McIntyre, D.C., Poulter, M.O., & Gilby, K. (2002). Kindling: some old and some new. *Epilepsy Research*, 50, 79-92.

Meeren, H.K.M., Pijn, J.P.M., Van Luijcklaar, E.L.J.M., Coenen, A.M.L., & Lopes da Silva, F.H. (2002). Cortical focus drives widespread corticothalamic networks during spontaneous absence seizures in rats. *Journal of Neuroscience*, 22, 1480-1495.

Michael, M., Holsinger, D., Ikeda-Douglas, C., Cammisuli, S., Ferbinteanu, J., DeSouza, C., et al. (1998). Development of spontaneous seizures over extended electrical kindling: Electrographic, behavioural and transfer kindling correlates. *Brain Research*, 793, 197-211.

Mormann, F., Andrzejak, R.G., Kreuz, T., Rieke, C., David, P., Elger, C.E., & Lehnertz, K. (2003b). Automated detection of a pre seizure state based on a decrease in synchronization in intracranial electroencephalogram recordings from epilepsy patients. *Physical Review E*, 67, 21912.

Mormann, F., Kreuz, T., Andrzejak, R.G., David, P., Lehnertz, K., & Elger, C.E. (2003). Epileptic seizures are preceded by a decrease in synchronization. *Epilepsy Research*, 53, 173-185.

Mormann, F., Lehnertz, K., David, P., & Elger, C.E. (2000). Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients. *Physica D*, 144, 358-369.

Moser, H.R., Weber, B., Wieser, H.G., & Meier, P.F. (1999). Electroencephalograms in epilepsy: analysis and seizure prediction within the framework of Lyapunov theory. *Physica D*, 130, 291-305.

Navarro, V., Martinerie, J., Le Van Quyen, M., Clemanceau, S., Adam, C., Baulac, M., et al. (2002). Seizure anticipation in human neocortical partial epilepsy. *Brain*, 125, 640-655.

Netoff, T., & Schiff, S.J. (2002). Decreased neuronal synchronization during experimental seizures. *Journal of Neuroscience*, 22, 7297-7307.

Nunez, P.L. (1995). Quantitative states of neocortex. In P.L. Nunez (Ed.), *Neocortical dynamics and human EEG rhythms (pp.3-67)*, New York: Oxford Press.

Pijn, J.P., Van Neerven, J., Noest, A., Lopez de Silva, F.H. (1991). Chaos or noise in EEG signals: dependence on state and brain site. *Electroencephalography and Clinical Neurophysiology*, 79, 371-81.

Racine, R.J., Steingart, M., & McIntyre, D. (1999). Development of kindling-prone and kindling-resistant rats: selective breeding and electrophysiological studies. *Epilepsy Research*, 35, 183-195.

Rajna, P., Clemens, B., Csibri, E., Dobos, E., Geregely, A., Gottschal, M., et al. (1997). Hungarian multi centre epidemiological study of the warning and initial symptoms (prodrome, aura) of epileptic seizures. *Seizure*, 6, 361-368.

Rogowski, Z., Gath, I., & Bental, E. (1981). On the prediction of epileptic seizures. *Biological Cybernetics*, 42, 9-15.

Rosenstein, M.T., Collins, J.J., De Luca, C.J. (1994). Reconstruction expansion as a geometry-based framework for choosing proper delay times. *Physica D*, 73, 82-98.

Rulkov, N.F., Sushchik, M.M., & Tsimring, L.S. (1995). Generalized synchronization of chaos in directionally coupled chaotic systems. *Physical Review E*, 51, 980-994.

Schreiber, T. (1999). Interdisciplinary application of nonlinear time series methods. *Physics Reports*, 308:1-64.

Schreiber, T., & Schmitz, A. (2000). Surrogate time series. *Physica D*, 142, 346-382.

Siegel, A.M., Grady, C.L., Mirsky, A.F. (1982). Prediction of spike-wave bursts in absence epilepsy by EEG power-spectrum signals. *Epilepsia*, 23, 47-60.

Steriade, M. (2001). *The intact and sliced brain*. Cambridge: The MIT Press.

Takens, F. (1981). Detecting strange attractors in turbulence. In: D. Rand & L.S. Young (Eds.), *Dynamical Systems and Turbulence (pp 366-381)*. New York: Springer.

Weinand, M.E., Carter, L.P., el-Saadany, W.F., Sioutos, P.J., Labiner, D.M., Oomen, K.J. (1997). Cerebral blood flow and temporal lobe epileptogenicity. *Journal of Neurosurgery*, 86, 226-32.

Appendix A

The data from the stimulated seizures during after discharge testing of the Fast and Slow strain rats is presented in the following pages. The animals were named T1 through T12, and as rat T10 died following surgery his data is not available. As each rat was tested for discharge threshold in the hippocampus and the amygdala, both of these seizures are presented, and each page is labeled with the animal's name, and the location of stimulation precipitating the seizure, in the manner

T11-RA(Fast),

which designates animal T11, stimulated in the right amygdala (RH would thus be the right hippocampus, and LA, LH the left amygdala and hippocampus respectively), and that this animal is from the Fast strain.

For each seizure there are three pages; the first page show the maxima of the divination function (Dv_{max}) vs the window number, thus the effect of time during seizure on the Dv_{max} . As there are twelve such series, one for each pair of signals, in each direction, they are presented in the same order each time. The graphs are thus

Graph 1, solid line = $Dv_{max}(LH,RA)$, dashed line = $Dv_{max}(RA,LH)$
 Graph 2, solid line = $Dv_{max}(LH,LA)$, dashed line = $Dv_{max}(LA,LH)$
 Graph 3, solid line = $Dv_{max}(LH,RH)$, dashed line = $Dv_{max}(RH,LH)$
 Graph 4, solid line = $Dv_{max}(RA,LA)$, dashed line = $Dv_{max}(LA,RA)$
 Graph 5, solid line = $Dv_{max}(RA,RH)$, dashed line = $Dv_{max}(RH,RA)$
 Graph 6, solid line = $Dv_{max}(LA,RH)$, dashed line = $Dv_{max}(RH,LA)$

And for each graph, the $Dv_{max} + 4.472$ standard deviations of an equal number of random points was plotted as a horizontal line; using the Tchebychev inequality this indicates that there is at most a 5% chance that a Dv_{max} above this value would be generated by chance.

The second page for each animal also has 6 graphs with a dashed and dotted line each; these graphs are the delay graphs, with the delay at which the Dv_{max} was found being plotted vs window number. They are presented in the same order as the previous graphs,

Graph 1, solid line = Δt (LH,RA), dashed line = Δt (RA,LH)
 Graph 2, solid line = Δt (LH,LA), dashed line = Δt (LA,LH)
 Graph 3, solid line = Δt (LH,RH), dashed line = Δt (RH,LH)
 Graph 4, solid line = Δt (RA,LA), dashed line = Δt (LA,RA)
 Graph 5, solid line = Δt (RA,RH), dashed line = Δt (RH,RA)
 Graph 6, solid line = Δt (LA,RH), dashed line = Δt (RH,LA)

This shows the lead-lag relationships over the course of the seizure.

The third page is simply the graphs of the actual signals (μV vs time), with time measured in milliseconds. As the windows are 3000ms long, and advance by 200ms shifts, we can see that any window number represents a 3000 ms stretch of the original series, and we can locate the represented segment of the signals by multiplying the window number by 200ms, and the Dv_{max} represents the 3 second segment following that time point. The graphs of the signals are labeled with the electrode locations, and are in order, the left hippocampus, right amygdala, left amygdala and right hippocampus.

T1-RA (Fast)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T1-RA (Fast)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T1-RA (Fast)

Left Hippocampus

Right Amygdala

Left Amygdala

Right Hippocampus

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T1-RH (Fast)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T1-RH (Fast)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T1-RH (Fast)

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T2-RA (Fast)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T2-RA (Fast)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T2-RA (Fast)

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T2-RH (Fast)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T2-RH (Fast)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T2-RH (Fast)

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T3-RA (Slow)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T3-RA (Slow)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T3-RA (Slow)

Left Hippocampus

Right Amygdala

Left Amygdala

Right Hippocampus

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T3-RH (Slow)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T3-RH (Slow)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T3-RH (Slow)

Left Hippocampus

Right Amygdala

Left Amygdala

Right Hippocampus

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T4-RA (Slow)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T4-RA (Slow)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T4-RA (Slow)

Left Hippocampus

Right Amygdala

Left Amygdala

Right Hippocampus

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T4-LH (Slow)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T4-LH (Slow)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T4-LH (Slow)

Left Hippocampus

Right Amygdala

Left Amygdala

Right Hippocampus

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T5-RA (Slow)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T5-RA (Slow)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T5-RA (Slow)

Left Hippocampus

Right Amygdala

Left Amygdala

Right Hippocampus

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T5-RH (Slow)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T5-RH (Slow)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T5-RH (Slow)

Left Hippocampus

Right Amygdala

Left Amygdala

Right Hippocampus

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T6-RA (Slow)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T6-RA (Slow)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T6-RA (Slow)

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T6-RH (Slow)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T6-RH (Slow)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T6-RH (Slow)

Left Hippocampus

Right Amygdala

Left Amygdala

Right Hippocampus

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T7-RA (Slow)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T7-RA (Slow)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T7-RA (Slow)

Left Hippocampus

Right Amygdala

Left Amygdala

Right Hippocampus

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T7-RH (Slow)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T7-RH (Slow)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T7-RH (Slow)

Left Hippocampus

Right Amygdala

Left Amygdala

Right Hippocampus

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T8-RA (Slow)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T8-RA (Slow)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T8-RA (Slow)

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T8-RH (Slow)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T8-RH (Slow)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T8-RH (Slow)

Left Hippocampus

Right Amygdala

Left Amygdala

Right Hippocampus

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T9-RA (Fast)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T9-RA (Fast)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T9-RA (Fast)

Left Hippocampus

Right Amygdala

Left Amygdala

Right Hippocampus

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T9-RH (Fast)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T9-RH (Fast)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T9-RH (Fast)

Left Hippocampus

Right Amygdala

Left Amygdala

Right Hippocampus

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T11-RA (Fast)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T11-RA (Fast)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T11-RA (Fast)

Left Hippocampus

Right Amygdala

Left Amygdala

Right Hippocampus

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T11-RH (Fast)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T11-RH (Fast)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T11-RH (Fast)

Left Hippocampus

Right Amygdala

Left Amygdala

Right Hippocampus

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T12-RA (Fast)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T12-RA (Fast)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T12-RA (Fast)

Left Hippocampus

Right Amygdala

Left Amygdala

Right Hippocampus

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

T12-RH (Fast)

Graph of $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed. The horizontal line is the $Dv_{\max}$ for a prediction of noise for the same number of embedded vectors.

T12-RH (Fast)

Graph of Δt associated with $Dv_{\max}$ vs. window number; each window is 3000 ms long and overlaps the previous by 2800 ms, thus each increment of a window indicates 200 ms have elapsed.

T12-RH (Fast)

Left Hippocampus

Right Amygdala

Left Amygdala

Right Hippocampus

Graphs of voltage (μV) vs. time (ms) for the seizure; thus 200 ms corresponds to advancing one window.

Appendix B

In testing the programs for both the calculation of the embedding dimension and the divination function it was necessary to generate data with known values; the Rossler and Lorenz attractors are two of the systems that were used in testing, and the equations for these systems and the parameters that were employed are listed below.

Rossler attractor:

$$\begin{aligned}\frac{1}{\tau_r} \frac{dr_x}{dt} &= -(r_y + r_z) - g_{lr} l_x \\ \frac{1}{\tau_r} \frac{dr_y}{dt} &= r_x + ar_y \\ \frac{1}{\tau_r} \frac{dr_z}{dt} &= b + r_x r_z + cr_z\end{aligned}$$

and the parameters were $a=0.2$, $b=0.2$, $c=5.7$. The term g_{lr} is a coupling term, which was set to 0 for modeling the Rossler attractor alone, but for systems involving a Rossler attractor coupled to a Lorenz attractor this variable determined the strength of the coupling, and l_x is the x variable for the Lorenz attractor. τ_r is a scaling constant used when we wish to set the scale of the attractor for the purpose of coupling; it is left at 1 otherwise.

Lorenz attractor

$$\begin{aligned}\frac{1}{\tau_l} \frac{dl_x}{dt} &= a(l_y - l_x) + g_{rl} r_x \\ \frac{1}{\tau_l} \frac{dl_y}{dt} &= bl_x - l_y - l_x l_z \\ \frac{1}{\tau_l} \frac{dl_z}{dt} &= -cl_z + l_x l_y\end{aligned}$$

and the parameters were $a=10$, $b=28$, $c=8/3$. The term g_{rl} is a coupling term, which was set to 0 for modeling the Lorenz attractor alone, but for systems involving a Lorenz attractor coupled to a Rossler attractor this variable determined the strength of the coupling, and r_x is the x variable from the Rossler attractor. τ_l is a scaling constant used when we wish to set the scale of the attractor for the purpose of coupling; it is left at 1 otherwise.

As the range of the two attractors is not guaranteed to be equal, a scaling was performed when the two systems were to be coupled; the values of the Rossler were scaled by setting $\tau_r=0.3$, and those of the Lorenz by setting $\tau_l=0.25$ in order to have similar ranges.