ornl

ORNL/TM-13215

OAK RIDGE
NATIONAL
LABORATORY

LOCKHEED MARTIN

Modeling and Analyzing Non-Seizure EEG Data for Patients with Epilepsy

W. F. Lawkins
N. E. Clapp, Jr.
C. S. Daw
L. M. Hively
V. Protopopescu
M. L. Eisenstadt

MANAGED AND OPERATED BY
LOCKHEED MARTIN ENERGY RESEARCH CORPORATION
FOR THE UNITED STATES
DEPARTMENT OF ENERGY

ORNL/TM-13215

Computer Science and Mathematics Division

Mathematical Sciences Section

MODELING AND ANALYZING NON-SEIZURE EEG DATA FOR PATIENTS WITH EPILEPSY

W. F. Lawkins [a]
N. E. Clapp,Jr. [a]
C. S. Daw [a]
L. M. Hively [a]
V. Protopopescu [a]
M. L. Eisenstadt [b]

[a] Oak Ridge National Laboratory
P.O. Box 2008, Bldg. 6012
Oak Ridge, TN 37831-6367
[b] Knoxville Neurology Clinic
930 Emerald Ave., Suite 815
Knoxville, TN 37917

Date Published: May 1996

Prepared by the
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831
managed by
Lockheed Martin Energy Research Corp.
for the
U.S. DEPARTMENT OF ENERGY
under Contract No. DE-AC05-96OR22464

Contents

List of Tables

List of Figures

Acknowledgements

We acknowledge members of the staff at the Knoxville Neurology Clinic, including Dr. Ronald E. Leppanen, Walter W. Holland, and Helen L. Leach, who provided expert assistance with the EEG data.

This research was partially sponsored by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory and by the Office of Basic Energy Research, U. S. Department of Energy.

MODELING AND ANALYZING NON-SEIZURE EEG DATA FOR PATIENTS WITH EPILEPSY

W. F. Lawkins
N. E. Clapp,Jr.
C. S. Daw
L. M. Hively
V. Protopopescu
M. L. Eisenstadt

Abstract

We present nonlinear analyses of non-seizure electroencephalogram (EEG) time series data from four epileptic patients. A non-seizure state is a period that is free of any part of an epileptic seizure, including the transition to a fully developed episode. EEG measurements are typically contaminated with a large amount of non-neurophysiological source information, generally called "artifact," which arises, for example, from eye movement, muscle tension, and physical motion. This study reflects three underlying objectives. The first objective is to gain some insight into how much variability in analysis results to expect from patients having similar clinical characteristics. The second objective is to investigate the impact of one specific type artifact, namely, eye movement, on the analysis results. A special feature of the research presented here is the introduction and testing of a filter for eye-movement artifact. Finally, the third objective is to determine if neurophysiological activity as viewed from two adjacent channels appears dynamically to be the same or different. Regarding the last objective, we also want to study the impact of artifact on observed coupling between channels.

1. Introduction

Two of the greatest challenges of neurological medicine are the accurate diagnosis and the effective treatment of epileptic seizures. Epilepsy remains one of the most prevelant, and yet poorly understood problems of medical science. The cost to the individual experiencing epilepsy and to society is truly enormous. The sine qua non for the evaluation of epileptic seizures is the electroencephalogram (EEG). The analog version has been a tool for many decades, and more recently, quantitative EEG using analyses of digitized data has become available. Despite the obvious advantages of quantitative EEG, it has offered relatively little to clinical epileptology. For instance, artifact confounds spike detection algorithms currently used in quantitative EEG and there are few quantitative markers of epileptic seizures. For this and a variety of other theoretical and practical reasons, there is cause for developing more sensitive and discriminating quantitative EEG methods than presently exist.

The purpose of this article is to present results of modeling and analyzing non-seizure EEG time series data from four patients with epilepsy. There is substantial evidence in the literature to support making the assumption that neurophysiological activity is a chaotic nonlinear dynamical process [12, 2, 15, 21, 3, 17], and the methods used to do the modeling and analyses presented here are based on that theory. Implementation of the methods we use depends on the fundamental reconstruction theorem due to Takens [20], which provides a means to represent such a system using time series data. Articles that review the theory and describe various measures of chaotic structure, as well as methods for calculating those measures, are Ruelle [18], Eckmann and Ruelle [7], Abarbanel et al. [1], and Grassberger et al. [11].

EEG measurements are typically contaminated with large amounts of artifact, which are data due to, for example, eye-movement, muscle tension, and physical motion. Because there is concern that artifact obscures the underlying brain dynamics, the usual procedure in a clinical setting is to eliminate the sections with artifact and analyze only the remaining artifact free data. However, typically 50% of a record is contaminated by artifact signals, and thus this procedure normally results in producing a relatively small file consisting of a number of short, distinct segments.

The primary purpose of this study is to gain some insight into how discriminating nonlinear techniques are compared to linear ones when applied to EEG data for epileptic patients. Beyond that, we also want to gain some insight on whether neurophysiological activity as observed by distinct channels appears to be the same or different and whether artifact interferes with that observed activity. If the modeling and analysis methods considered here are ultimately shown to be transparent to the presence of artifact, the current clinical practice of explicitly removing those segments may prove to be unnecessary. Another potential means for coping with artifact may be to filter it. A band-pass filter is introduced that can be tuned to different types of artifact, and we make a preliminary assessment of its effectiveness.

The analysis is organized into three parts to reflect the primary and secondary objectives described above. The first part considers variability in analysis results between patients having similar clinical characteristics. Here, we mimic the clinical setting by first removing artifact contaminated segments from the record. The second part consid-

ers whether neurophysiological activity as viewed from two adjacent channels appears dynamically to be the same or different and, in addition, considers the impact of one specific type artifact, namely, eye movement, on that appearance. The performance of the artifact filter is considered here, too. Finally, the third part considers coupling between adjacent channels. Again, we consider the impact of artifact and the artifact filter on that observed coupling.

As described above, the motivation for the research presented here is the goal to develop better methods for analyzing EEG data. The methods we use consist, first, of a technique for constructing a model of a nonlinear dynamical process from time series data and, second, of methods for estimating quantitative measures of the ergodic structure of the resulting model. In order to make clear what is meant by a model, how to construct a model, and how to analyze a model, before presenting the three part results described above we go through the procedure step by step, illustrating each step using one of the EEG files.

We described the EEG records above as being non-seizure data. To be more exact, each record covers a tranquil period that does not contain a seizure, the transition period leading to a seizure, and the recovery period following a seizure. One EEG record from each of four patients is used for this study. Each of those records consists of sixteen channels arranged according to the longitudinal bipolar montage EEG protocol, which is illustrated in Fig. 1. Further, each record covers a single continuous measurement period of 600 seconds. The data was obtained using instrumentation with parameter values

$$\begin{aligned} f_s &= \text{sampling rate} \\ &= 512\,Hz\,, \\ f_c &= \text{low-pass filter cutoff} \\ &= 100\,Hz\,, \\ f_h &= \text{high-pass filter cutoff} \\ &= 1.5\,Hz\,. \end{aligned} \tag{1}$$

We want to describe in a generic way how an EEG record is used to generate data files used in this study. Selecting a channel from the EEG record, a file is created from the given record that consists of time series segments that are artifact free. A second file is created that contains time series segments with eye movement artifact such that those segments do not overlap in time any part of the segments in the first file. The first file is used to analyze the normal, artifact free process as observed by the given channel and is intended to mimic how analysis is done in the clinical setting. The second file concatenated with the first is used to represent the normal process with the addition of eye movement artifact.

Referring to Fig. 1, note that Channel 13 is one of two channels closest to an eye, and that Channel 14 is once removed from Channel 13 along the hemispheric boundary. We expect Channel 13 to be at least as affected by eye-movement artifact as any channel. Further, if the analysis techniques used here can detect coupling between channels, we expect to observe that between Channels 13 and 14. Also, if there is decay in the influence of eye-mevement artifact with distance from an eye, we expect to see that using those two channels. For those reasons, we use data from Channels 13 and 14 in this study.

Table 1 is a list of the files constructed from the four EEG records used here. Again, the files contain artifact-free segments and segments with only eye-movement artifact. The extensions 13 and 14 refer to Channels 13 and 14, respectively. The letters "em" are used to designate eye-movement artifact data files. Note that there are two such files, both for Patient 1. Further, note that there is one artifact-free file for each patient, all derived from Channel 13. We note two additional facts regarding the arrangement of segments within a file. First, within a given file, the ordering of segments reflects increasing time relative to the EEG record. Second, in those cases where there are pairs of files for Channels 13 and 14, like $fe1ch13$ and $fe1ch14$, the corresponding segments are aligned in time, namely, the first segment from $fe1ch13$ is aligned in time with the first segment from $fe1ch14$, and so forth.

Patient	File Name	Channel	File Description	No. Segments
1	fe1ch13	13	artifact-free	4
1	fe1ch14	14	artifact-free	4
1	fe1em13	13	eye-movement artifact	4
1	fe1em14	14	eye-movement artifact	4
2	25402c13	13	artifact-free	4
3	28049c13	13	artifact-free	5
4	28158c13	13	artifact-free	4

Table 1: A list of the files constructed from the four EEG records.

The remainder of this paper is presented in four sections. Section 2 reviews mathematical tools used for time series analysis and for constructing an approximation of the nonlinear dynamical process represented by the EEG data. Further, measures of the ergodic structure of chaotic processes that we estimate and present in this paper are defined. At each step an example is given using one of the files listed in Table 1. Section 3 describes a quadratic filter designed to remove low-frequency artifact corresponding to eye movement. Section 4 reports the results of the nonlinear analyses organized according to the three parts outlined above. First, we examine the nonlinear variability among all four patients for one channel of artifact-free data. Second, using data from Channels 13 and 14 for Patient 1, we apply the artifact filter described in Section 3 to the with-artifact data to generate artifact filtered data. We then compare nonlinear models and analyses using artifact-free, with-artifact, and artifact-filtered data from the two channels. Third, we calculate a nonlinear measure of coupling between Channels 13 and 14 for Patient 1 using artifact-free, with-artifact, and artifact-filtered data and assess the differences. Finally, Section 5 presents our conclusions.

2. Nonlinear Processes and Time Series

We begin this section by reviewing the definition of a nonlinear dynamical process and describing the theory that provides for constructing an approximate model of such a process using time series. Next, we review some mathematical tools used to determine fundamental information about the process from the time series which is needed to construct the model of the nonlinear process. We then describe the technique used

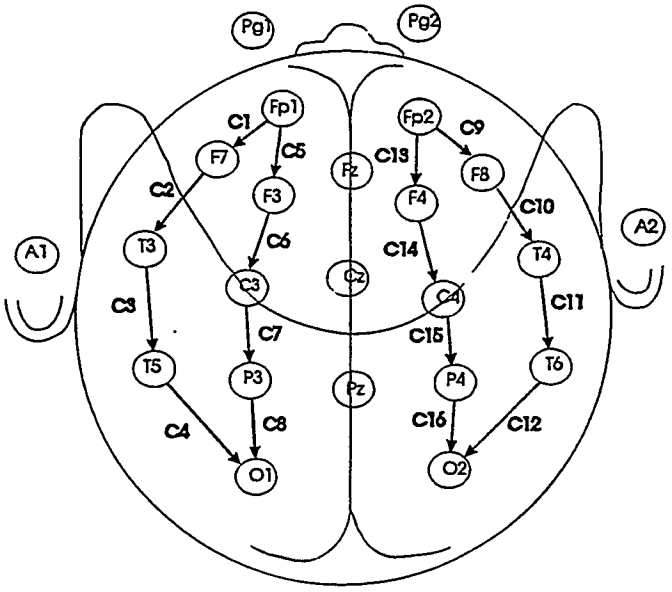

Figure 1: Bipolar montage used to obtain EEG measurements.

for determining the model. Finally, we define and describe the quantitative ergodic measures of the reconstructed nonlinear process presented in this paper. We illustrate each step using the file $fe1ch13$ from Table 1, which is artifact-free data from Channel 13 for Patient 1.

2.1. Nonlinear Dynamical Processes

Using $E^{\bar{n}}$ to denote $\bar{n}$-dimensional Euclidean space, let

$$f : E^{\bar{n}} \to E^{\bar{n}}$$

be a diffeomorphism. Suppose $M \subset E^{\bar{n}}$ is a compact, $\bar{n}$-dimensional differentiable manifold and that f restricted to M is a diffeomorphism of M. Let A be a compact subset of M such that f maps A onto A. Further, let $U \supset A$ be an open subset of M such that

$$\lim_{i \to \infty} f^i(U) = A.$$

Finally, let p be an ergodic probability measure on A with respect to the transformation f, which is to say that, if $V \subset A$ is measurable with respect to p, then $p(f(V)) = p(V)$.

We refer to $E^{\bar{n}}$ as "state space" and to A as an "attractor." The point of view taken here is that nonlinear dynamical processes, including chaotic systems, are represented by triples of the form (A, f, p), that is, an attractor in state space, a diffeomorphic mapping of that attractor onto itself, and an ergodic probability measure with respect to that attractor and diffeomorphic mapping.

Suppose $a \in A$ is the state of the experimental system at the instant we begin to

observe it. Then, $\{f^i(a)\}_{i=0}^{\infty}$ is a time series of states visited by the process. Now, suppose $y : A \to \mathbb{R}$ is a real-valued function, or observable, on A. Then,

$$y_i = y(a_i) = y(f^i(a)), \quad i = 0, 1, 2, \ldots,$$

is a real-valued time series of the experimental system.

The following is a theorem due to Takens[20] that constitutes the foundation for the time series analysis methods described here. As used in the statement of the theorem, "smooth" means at least C^2.

Theorem 1. *(Takens) Let $M \subset E^{\bar{n}}$ be a compact, differentiable manifold of dimension $\bar{n}$. Further, let $f : M \to M$ be a smooth diffeomorphism of M and let $y : M \to \mathbb{R}$ be a smooth, real-valued function on M. It is a generic property that the map*

$$\Phi : M \to E^{2\bar{n}+1}$$

defined by

$$\Phi(a) = (y(a), y(f^1(a)), \ldots, y(f^{2\bar{n}}(a))^T, \quad a \in M,$$

where T denotes matrix transpose, is a smooth embedding.

For some choice of values for the positive integers (n, k), consider the object $\tilde{A}_y \subset E_y^n$, where

$$\tilde{A}_y = \{\mathbf{y}_i\}_{i=0}^{\infty} \tag{2}$$

and $\mathbf{y}_i$ is defined by

$$\mathbf{y}_i = (y_i, y_{i+k}, y_{i+2k}, \ldots, y_{i+(n-1)k})^T. \tag{3}$$

Provided $n \geq 2\bar{n} + 1$, Takens' reconstruction theorem implies the mapping

$$\tilde{\Phi} : A \to \tilde{A}_y$$

defined by

$$\tilde{\Phi}(a_i) = \mathbf{y}_i, \quad i = 0, 1, 2, \ldots,$$

produces a smooth embedding $\tilde{\Phi} : A \to E_y^n$. Define

$$\begin{aligned} \tilde{f}_y &= \tilde{\Phi} \circ f \circ \tilde{\Phi}^{-1} \\ \tilde{p}_y &= p \circ \tilde{\Phi}^{-1}. \end{aligned}$$

Then, the triple $(\tilde{A}_y, \tilde{f}_y, \tilde{p}_y)$ constitutes a faithful representation of the dynamical system (A, f, p).

The pair of integers (n, k) used to define the vector $\mathbf{y}_i$ Eq. (3) are referred to as the embedding dimension and time lag, respectively. Choosing appropriate values for (n, k) is a primary objective for an analysis leading to a model $(\tilde{A}_y, \tilde{f}_y, \tilde{p}_y)$ of a nonlinear dynamical process.

2.2. Low-Pass Filters

Low-pass filtering plays an important role in conditioning time series data. The significance of such conditioning for constructing empirical nonlinear dynamical models is discussed later. Here we describe the specific low-pass filter used for the analyses presented in this article.

Consider the first-order, linear low-pass filter

$$\frac{1}{\omega}\frac{dy}{dt} + y = x,$$

where x is input, y is output, $\omega = 2\pi f_c$, and f_c is the cutoff frequency. Integrating this equation over the interval $[t_i, t_i + t_s]$ produces

$$\begin{aligned} y_{i+1} &= y_i \exp(-\omega t_s) + x_i[1 - \exp(-\omega t_s)], \\ &= a y_i + b x_i, \end{aligned} \tag{4}$$

where $x_i, i = 0, 1, 2, ...$ is a time series obtained from $x(t)$ using the sampling time t_s. Eq. (4) is referred to in control theory as a first-order lag, or first-order infinite impulse response, low-pass filter and is frequently used to simulate the behavior of measurement instuments[16]. We use the fourth-order filter that results by applying the first-order filter Eq. (4) in series four times, which is to say output from the first stage is input to the second, output from the second stage is input to the third, and so forth.

Figure 2 displays a segment of the EEG data before and after applying the low-pass filter. The filter cutoff values displayed in Fig. 2 are normalized by the sampling frequency. The normalized filter cutoff value for Fig. 2a is $f_c = .5$, which is half the sampling frequency. That value for the cutoff frequency is commonly known as the "Nyquist" frequency and is used in measurement instrumentation to prevent "aliasing." The sampling time, sampling frequency, and the dimensional value of f_c shown in Fig. 2b are, respectively,

$$t_s = \tfrac{1}{512} s, \quad f_s = 512\, Hz, \quad f_c = 50\, Hz. \tag{5}$$

2.3. Power Spectra and Autocovariance

Power specta analysis of discrete time series is well established and we refer to Blackman and Tukey [5] for detailed information on that topic. Here, we want to briefly review the relationship between the power spectrum and autocovariance function of a time series $y = \{y_i\}_{i=1}^N$. The autocovariance estimate for lag k is

$$c_k = \frac{1}{N}\sum_{i=1}^{N-k}(y_i - \bar{y})(y_{i+k} - \bar{y}) \ , \ k = 0, 1, ..., N-1 \ , \tag{6}$$

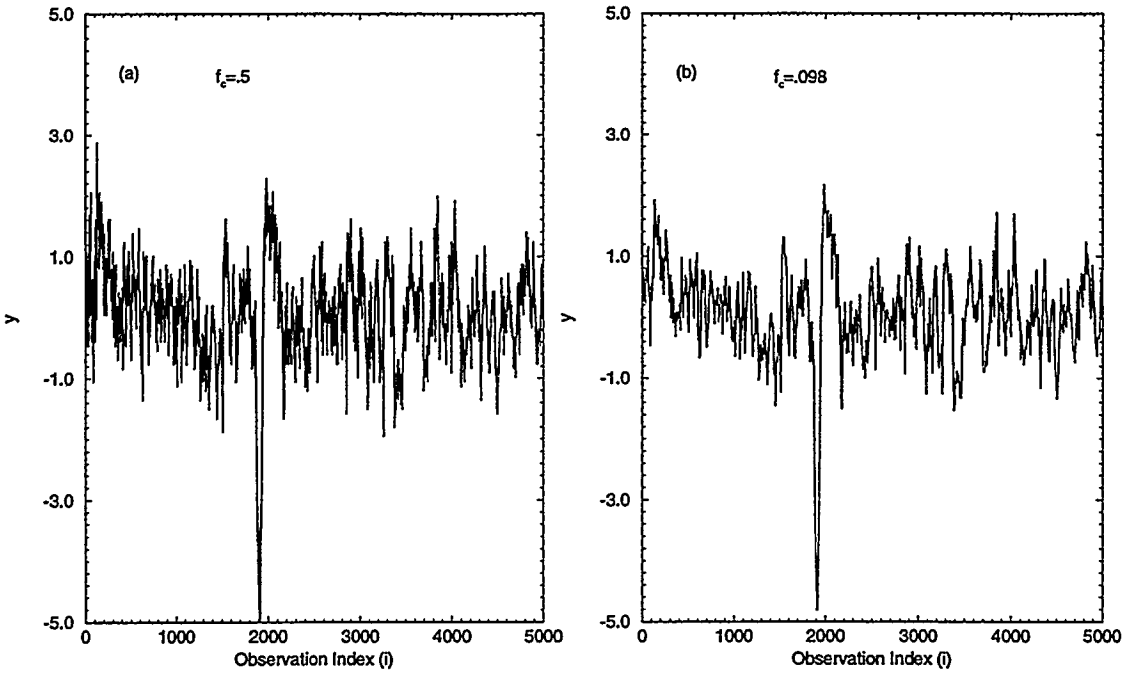

Figure 2: Example segments for the EEG data, y, as a function of observation index, i: (a) as-measured data ($f_c = .5$); (b) conditioned EEG data ($f_c = .098$).

where $\bar{y}$ is the mean value of y. Suppose N is odd. Then we can write $N = 2q + 1$. The time series y can be decomposed into Fourier modes according to the equations

$$y_i = \alpha_0 + \sum_{j=1}^{q} (\alpha_j \cos(2\pi f_j i) + \beta_j \sin(2\pi f_j i)) \, , \; i = 1, ..., N, \tag{7}$$

where $f_j = j/N$ and

$$\begin{aligned} \alpha_0 &= \bar{y} \, , \\ \alpha_j &= \frac{2}{N} \sum_{i=1}^{N} y_i \cos(2\pi f_j i) \, , \\ \beta_j &= \frac{2}{N} \sum_{i=1}^{N} y_i \sin(2\pi f_j i) \, . \end{aligned} \tag{8}$$

The power spectrum for the time series y is

$$P(f_j) = \frac{N}{2} (\alpha_j^2 + \beta_j^2) \, , \; j = 1, 2, ..., q. \tag{9}$$

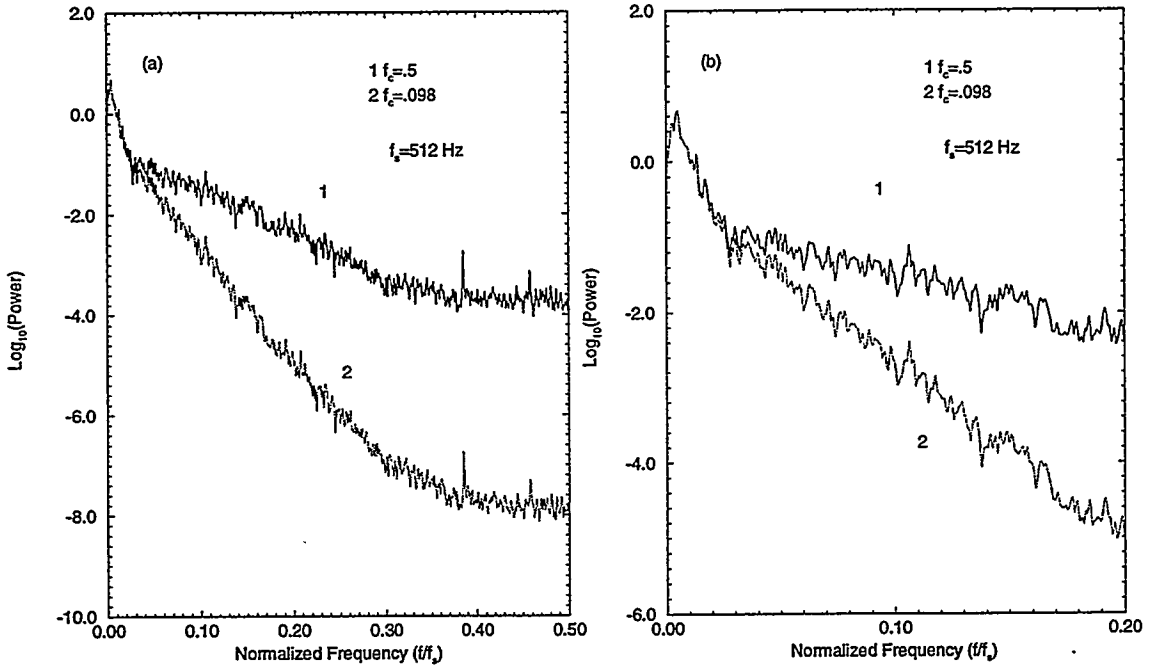

Figure 3: Power spectra, $\mathrm{Log}_{10}(\mathrm{Power})$, as a function of normalized frequency, f/f_s, for EEG data: (a) EEG data; (b) magnified view of (a).

Finally, the sample spectrum $P(f)$ can be calculated from the sample autocovariance by the cosine transformation

$$P(f_j) = 2[c_0 + 2\sum_{k=1}^{N-1} c_k \cos(2\pi f_j k)] \ , \ f_j = \frac{j}{N} \ . \tag{10}$$

In general, most of the variation in the autocovariance function occurs for relatively small values of lag k, so that it follows most of the energy in the power spectrum $P(f)$ for high frequencies derives from the autocovariance function for small values of k. We reference this observation below.

Figure 3 displays power spectra for the EEG data, both for the as-measured data and for the conditioned time series obtained using the low-pass filter described in the previous section. The choice of a low-pass filter cutoff value is guided by the desire to preserve power in the dominant low frequency band and to minimize power above that band. The magnified frame in Fig. 3 supports the view that those conditions are achieved by the choice made for the filter cutoff value.

2.4. Mutual Information

Mutual information is a nonlinear measure of the extent to which one random variable is a function of another [19]. Let (x, y) be an $\mathbb{R}^2$-valued random variable. Further, let $\rho(x, y)$ be the joint probability density of (x, y), and let $\rho(x), \rho(y)$ be the probability densities of x, y, respectively. Then, the mutual information of the random variables

x, y is, by definition,

$$M(x,y) = \int_{\mathbb{R}^2} m(x,y)dxdy, \tag{11}$$

where

$$m(x,y) = \rho(x,y)ln\frac{\rho(x,y)}{\rho(x)\rho(y)}.$$

We note several properties of mutual information M. First, by definition, mutual information is symmetric in its arguments. Second, $M(x,y) \geq 0$ and $M(x,y) = 0$ if and only if x, y are independent. Third, if y is a function of x, then $M(x,y)$ is unbounded.

Consider now the time-series $\{y_i\}_{i=1}^N$, and suppose y is a measurement from a stationary stochastic process. For an arbitrarily fixed value of the time delay k, we define the bivariate $\mathbb{R}^2$-valued random variable (y_i, y_{i+k}) and the mutual information function associated with that time series by

$$M(k) = M(y_i, y_{i+k}). \tag{12}$$

The function $M(k)$ is a nonlinear measure of the dependence of two observations from the time series y separated by the time delay k. In general, mutual information is regarded as a more appropriate measure of independence versus dependence for a nonlinear process than the autocovariance function Eq. (6) [8, 9]. Like the autocovariance function, $M(k)$ in general varies most rapidly for small values of delay k, and, referring to the equation expressing the relationship between the autocovariance and power spectrum Eq. (10) and the observation made following that equation, the time scales associated with the most rapidly varying segment of mutual information translate to relatively high frequencies that are not resolved in the nonlinear process reconstruction described below.

Figure 4 displays mutual information for the EEG data, both for the as-measured and conditioned data sets. Note the difference in the two curves. Mutual information for the as-measured data, which has more noise than the conditioned data, falls off more rapidly than that for the conditioned file. For both data sets, the variation in mutual information is small for lag values $k > 20$, which corresponds to a normalized frequency

$$f \approx 1/20 = .05\,.$$

Referring to the power spectra for the as-measured and conditioned data Fig. 3, note that the conditioned data loses power rapidly for normalized frequencies $f > .05$, thus magnifying the relative power for frequencies $f < .05$. The low-pass filter cutoff value used to condition the data, $f_c = .098$, isolates the dominant frequency content of the data, which is reflected in Fig. 3.

2.5. Process Model

We review a technique, which is based on a procedure introduced by Broomhead and King [6], for constructing a model of a nonlinear process. The technique described here is a refinement of that given by Lawkins et al. [13]. The refinement concerns

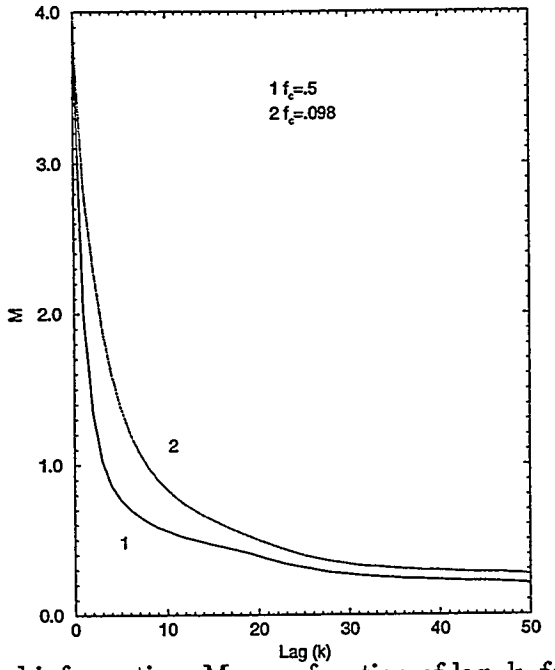

Figure 4: Mutual information, M, as a function of lag, k, for the EEG data.

the combined use of time domain information represented by mutual information with frequency domain information represented by power spectra to construct a model. While the use of mutual information for constructing models was introduced by Fraser and Swinney [8], the way that information is used here is different than that described by those researchers and practiced, for example, by Abarbanel et al. [1].

Recall that the primary task is to determine appropriate values for the reconstruction parameters (n, k), where n is the embedding dimension and k is the time delay. Given values for those parameters, we have

$$\tilde{A}_y = \{\mathbf{y}_i\}_{i=1}^{N} \subset E_y^n \,, \tag{13}$$

where

$$\mathbf{y}_i = (y_i, y_{i+k}, y_{i+2k}, \cdots, y_{i+(n-1)k})^T \,. \tag{14}$$

Given the parameter values (n, k) and the sampling time t_s, we define

$$\begin{aligned} t_w &= (n-1) \times k \times t_s \\ &= \text{the window time scale} \,, \\ f_w &= t_w^{-1} \\ &= \text{the window frequency} \,. \end{aligned} \tag{15}$$

Next, we describe an orthonormal transformation

$$\Psi : E_y^n \to E_\theta^n$$

based on singular-value analysis. Define the correlation matrix

$$M = \frac{1}{N}\sum_{i=0}^{N} \mathbf{y}_i \mathbf{y}_i^T .$$

The eigenvalue/eigenvector pairs of the matrix M are the singular values and principal modes associated with the trajectory, or multivariate time series, $\{\mathbf{y}_i\}_{i=1}^N$. We denote the singular values and principal modes in pairs by

$$\{(\sigma_j^2, \psi_j)\}_{j=1}^n, \tag{16}$$

where ordering is determined by

$$\sigma_1^2 \geq \sigma_2^2 \geq ... \geq \sigma_n^2. \tag{17}$$

Define the $n \times n$ matrix Ψ by

$$\Psi = (\psi_1, \psi_2, ..., \psi_n) ,$$

so that the principal mode ψ_j is the j-th column of Ψ. Then

$$\theta = \Psi^T \mathbf{y} \tag{18}$$

defines an orthonormal transformation $\Psi : E_y^n \rightarrow E_\theta^n$ that transforms

$$\tilde{A}_y = \{\mathbf{y}_i\}_{i=1}^N \quad \overset{\Psi}{\rightarrow} \quad \tilde{A}_\theta = \{\theta_i\}_{i=1}^N .$$

We want to make two observations regarding the transformation Eq. (18). The first is that, in general, the principal mode ψ_j is expected to have a total of j relative maxima and minima. Thus we expect the j-th mode to have approximately $j/2$ waves over the time span of a point $\mathbf{y}_i$, which is the window time scale t_w Eq. (15). Consequently, the j-th coordinate in E_θ^n, which is the inner product of $\mathbf{y}_i$ with ψ_j, corresponds approximately to information in the time series resolved by the frequency

$$f_j = [\frac{t_w}{j/2}]^{-1} = j \times \frac{1}{2} f_w , \tag{19}$$

where f_w is the window frequency Eq. (15). The second observation is that the singular values Eq. (16) are estimates of the second moments of each coordinate value in the set $\tilde{A}_\theta$ and, as such, determine length scales in phase space E_θ^n for $\tilde{A}_\theta$.

We select candidate values for the reconstruction parameters (n, k) based primarily on the following two conditions. First, referring to Eq. (19), we want the lowest order, principal component frequencies f_j, $j = 1, \ldots, m$, where $m \leq n$, to resolve information in the time series, as represented by the power spectrum, considered to be most important. Second, we want the corresponding distribution of singular values, σ_j^2, $j = 1, \ldots, m$, Eq. (16), to resolve the state-space range of length scales associated

with the frequencies f_j, $j = 1, \ldots, m$. Recall the ordering of singular values Eq. (17). If σ_{m+1} is sufficiently small, the projection

$$P^m : E_\theta^n \to E_\theta^m \,, \tag{20}$$

where E_θ^m is the linear subspace of E_θ^n spanned by to the first m principal components, will result only in eliminating small scale, noisy detail in the highest order coordinates. We refer to the resulting reconstruction by the ordered triple of parameters (n, k, m).

Figure 5 displays singular values and normalized coordinate frequencies for the three pairs of candidate model parameter values

$$(n, k) = (6, 3), (9, 3), (12, 3) \,. \tag{21}$$

The normalized window time scales and frequencies Eq. (15) for those parameter values are, respectively,

$$\begin{array}{lcccc} (n-1) \times k & = & 15\,, & 24\,, & 33\,, \\ f_w \times t_s & = & .67(-1)\,, & .42(-1)\,, & .30(-1)\,. \end{array} \tag{22}$$

Referring to mutual information Fig. 4, the lag value $k = 3$ accounts for the most rapidly varying segment of the mutual information curve and the last two normalized window time scales cover the range of exponential variation. Referring to the power spectrum Fig. 3, note that the last two normalized window frequencies are within the low frequency band isolated by the low-pass filter. We delay until the following section making a judgement as to the adequacy of the three candidate models.

In general, the low-pass filter described in Section 2.2 is used to reduce the frequency content of the time series for frequencies greater than the upper bound of the frequency band in the power spectrum considered to include significant information about the process of interest. Recall from Section 2.4 that the extent of the most rapidly varying segment of mutual information also provides an estimate of the upper bound of that frequency band. The impact of such low-pass filtering is to reduce the potential noise-like effect of high frequencies on the coordinate frequencies Eq. (19) associated with a model and, hence, on the analyses that depend on that model.

2.6. Correlation Dimension

We review the definition of correlation dimension. Let $\Omega = \{\omega_i\}$ be a partition of the attractor A. For an arbitrarily fixed real number $q > 0$, we define

$$C_q(\Omega) = \{\sum_i p(\omega_i) p^{q-1}(\omega_i)\}^{\frac{1}{q-1}}$$

and the corresponding q-th order information associated with Ω by

$$I_q(\Omega) = -ln(C_q(\Omega)).$$

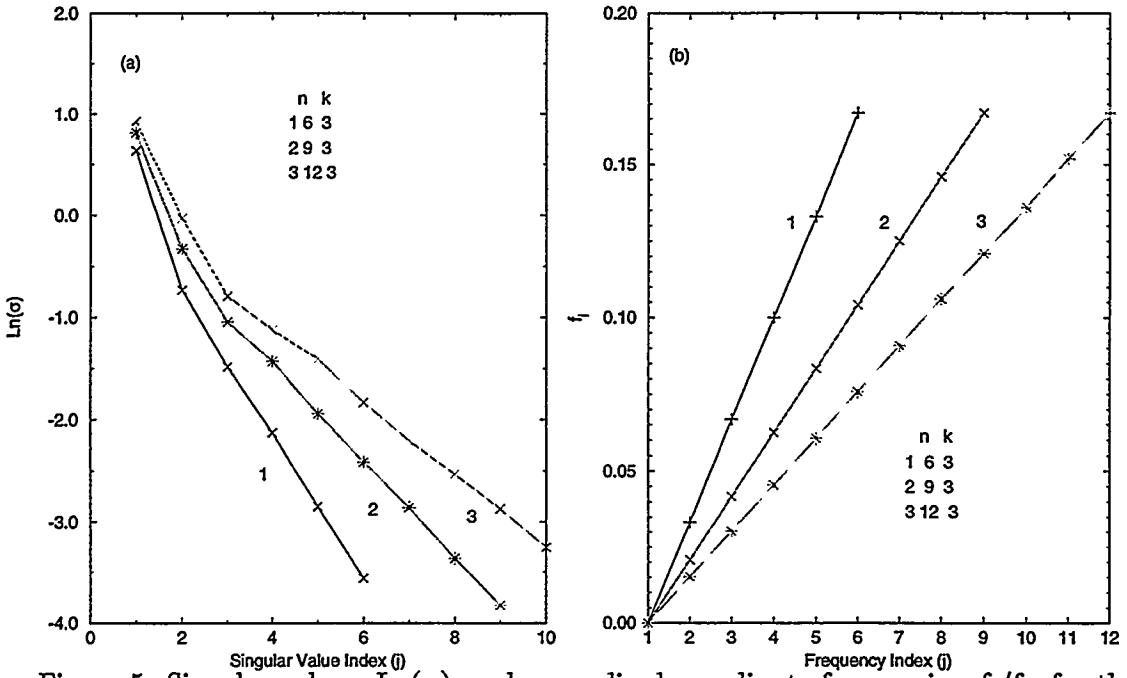

Figure 5: Singular values, $\mathrm{Ln}(\sigma)$, and normalized coordinate frequencies, f_j/f_s, for the EEG data: (a) singular values; (b) coordinate frequencies.

If Ω is a partition of length-scale ϵ, we also use the notation

$$C_q(\epsilon) = C_q(\Omega)$$

and

$$I_q(\epsilon) = -ln(C_q(\epsilon)). \tag{23}$$

If

$$C_q(\epsilon) \propto \epsilon^{\nu_q} \tag{24}$$

over some range of length-scale ϵ, then the process has q-th order "fractal structure" of dimension ν_q over that range of ϵ. The "correlation" dimension results if $q = 2$, which corresponds to the analysis results presented in this paper. Viewed as a function of ϵ, $I_2(\epsilon)$ is referred to as the "correlation integral." Henceforth, we use the notation

$$\begin{aligned} \nu &= \nu_2 , \\ C &= C_2 , \\ I &= I_2 . \end{aligned}$$

Suppose Ω is a partition of length scale ϵ. Then $I(\epsilon)$ is a measure of the expected number of base-e digits two points drawn from the same member of Ω are expected to have in agreement in each coordinate.

The method used to calculate the correlation integral, and hence the correlation dimension, is due to Grassberger and Procaccia [10]. Fig. 6 displays correlation integrals corresponding to the (n, k) parameter values used in Fig. 5. According to Eq. (24), the correlation dimension is the slope of the curve $-I(ln(\epsilon))$, which appears to be converged

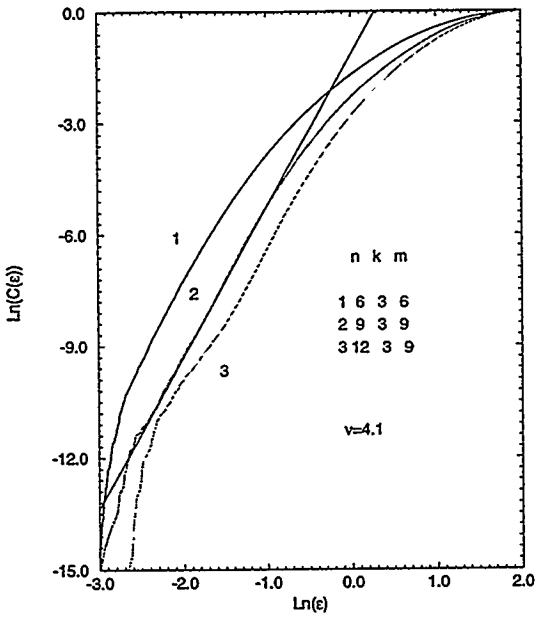

Figure 6: Correlation integrals, Ln(C(ϵ)), as a function of the partition, Ln(ϵ), for EEG data using (n,k) parameter values displayed in Figure 5.

for the candidate models Eq. (21)

$$(n,k,m) = (9,3,9), \tag{25}$$
$$(n,k,m) = (12,3,9), \tag{26}$$

yielding the value

$$\nu = 4.1\,. \tag{27}$$

Note the difference in state-space length scales for which fractal structure is resolved by the two models. Referring to the normalized window frequencies Eq. (22) and singular values Fig. 16, the effect of moving the window frequency from $f_w = .42(-1)$ to $f_w = .30(-1)$ is to translate resolved state-space structure to larger length scales.

We use the correlation integral results to judge the adequacy of a candidate model, and in this instance we make the judgement that the models Eqs. (25), (26) are both adequate. Because fractal structure is realized over a greater range of information for the model $(n,k,m) = (9,3,9)$, for subsequent analyses we use that model.

2.7. Correlation Entropy

We refer to Ben-Mizrachi et al. [4]. for the definition of entropy and for the idea used to calculate it. We note that a modified version of the method presented by those researchers is used to make calculations given in this article.

We review the definition of correlation entropy. First, suppose Ω^0, Ω^1 are partitions of A. We say Ω^1 is a refinement of Ω^0 if every member of Ω^0 can be represented as a union of members from Ω^1. For example, let Ω^0, Ω be partitions of A and define the

refinement Ω^1 of Ω^0 using Ω by

$$\begin{aligned} \Omega^1 &= \Omega^0 \vee \Omega \\ &= \{\omega_0 \cap \omega : \omega_0 \in \Omega^0, \omega \in \Omega\}. \end{aligned} \tag{28}$$

Given Ω^0, a partition of A, we are going to define a sequence $\{\Omega^i\}$ of refinements of Ω^0 using the nonlinear dynamical process f. Define the partition Ω^{-i} by

$$\Omega^{-i} = \{f^{-i}(\omega_0) : \omega_0 \in \Omega^0\}. \tag{29}$$

We recursively define

$$\Omega^{i+1} = \Omega^i \vee \Omega^{-(i+1)}. \tag{30}$$

The important point to note regarding the definition of the partition Ω^i is that, if two points, $\mathbf{u}$, $\mathbf{v}$, belong to the same member of Ω^i, then the i-th iterates $f^i(\mathbf{u})$, $f^i(\mathbf{v})$ belong to the same member of the initial partition Ω^0.

Suppose Ω^0 is of length-scale ϵ. Then, for ϵ sufficiently small and i sufficiently large, the q-th order entropy of the system is

$$K_q \simeq \frac{1}{i} I_q(\Omega^i) = \frac{1}{i} I_q^i. \tag{31}$$

It is the asymptotic rate at which information increases with time step for ϵ sufficiently small. In the case $q = 2$, K_2 is called the correlation entropy. From this point we use the notation

$$K = K_2 .$$

Entropy is an integral measure of expected temporal predictability. For a chaotic process, trajectories initiated from two nearby points on the attractor A are expected to diverge exponentially. Recall the observation made following Eq. (30). For a given error in predictability, ϵ, entropy is the asymptotic rate of increase in the number of base-e digits two initial points on A are expected to have in agreement if we observe that the trajectories initiated from those points continue to remain within ϵ of one another as time increases.

Consider Fig. 7. We note that the correlation integral for the reconstruction parameter values $(n, k, m) = (9, 3, 9)$, which is illustrated in Fig. 6, corresponds to the initial member of the family of integrals displayed in Fig. 7a. For a fixed value of $ln(\epsilon)$, Fig. 7a illustrates how information I^i is increasing with time step, i, and that there is a range of $ln(\epsilon)$ over which the slope of I^i is invariant with i. Further, using the fixed value $I = 4.0$, Fig. 7b displays how $ln(\epsilon)$ is increasing with i and that there is a scale for i where the rate of increase is constant. The correlation entropy, K, is approximated as a linear rate of increase of I^i. To estimate K, first, we determine an appropriate time scale, $\bar{i}$. Using that time scale, next we estimate the constant slope of the family I^i, denoted s, which is referred to as the "entropy dimension," and then we estimate the linear rate of increase of $ln(\epsilon)$, denoted ω, which is called the "characteristic wave

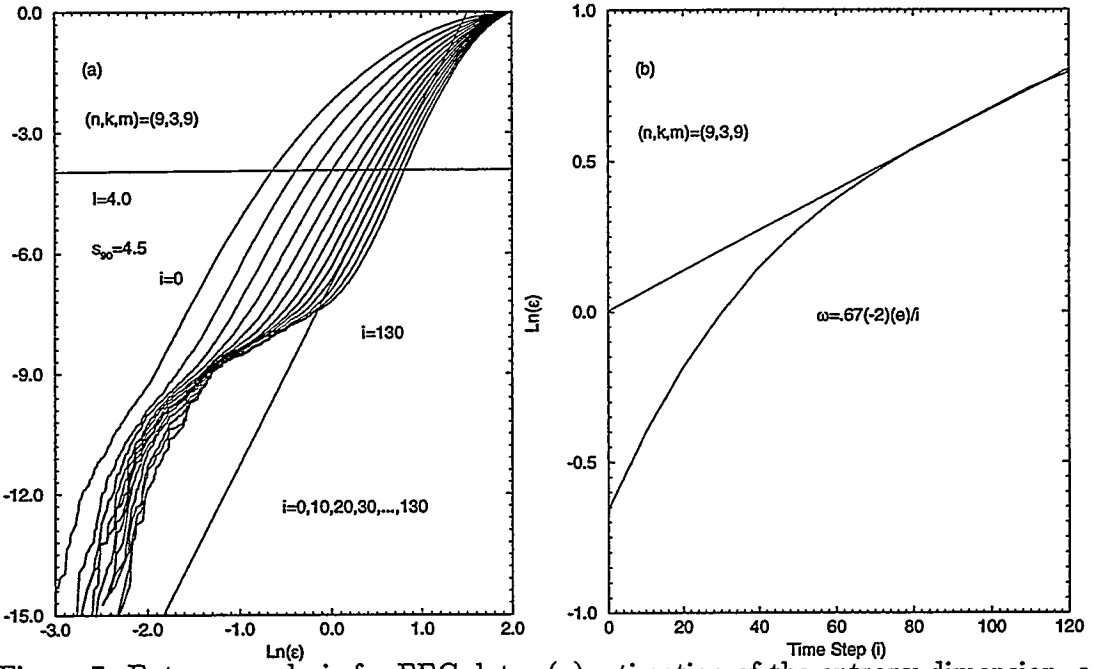

Figure 7: Entropy analysis for EEG data: (a) estimation of the entropy dimension, s, as the slope of $-\mathrm{I}^{90}(\mathrm{Ln}(\epsilon))$; (b) estimation of the characteristic wave speed, (ω), as the constant rate of change of $\mathrm{Ln}(\epsilon)$ with i using the relation $4.0 = \mathrm{I}^{\mathrm{i}}(\mathrm{Ln}(\epsilon))$.

speed." Entropy is estimated as $K = s \times \omega$. Referring to Fig. 7, we have

$$\begin{aligned} s &= 4.5 \quad , \\ \omega &= .67(-2)\,(e)/i \quad , \\ K &= .31(-1)\,(e)/i \quad , \end{aligned} \tag{32}$$

where the the time scale is

$$\bar{i} = 72\,i\,.$$

The rationale used to analyze $\bar{i}$ is described in the following section.

Consider the question of variability in estimates of K. Referring to Fig. 7, note that the integrals I^i appear to be linear and uniformly spaced for information values $3 \leq I \leq 7$ and time steps $80 \leq i \leq 130$. In fact, using the representation $K = s \times \omega$, the variation in estimates of K is relatively small over that range of I.

Concerning units, if

$$|u_i - v_i| \propto e^{\lambda i},\ i = 0, 1, 2, \ldots,$$

where u_i, v_i are trajectories in state space, we say the trajectories are diverging at the exponential rate λ and refer to the units as units (λ)=(e)/i. Note that those are also the units of ω and K Eq. (32).

2.8. Poincaré Return Map Analysis

We use a Poincaré return map approximation to estimate the dominant cyclic time-scale of the process and to estimate a Lyapunov exponent[7]. The description here is algorithmic.

We work with the initial embedding, that is, with the reconstruction in E_y^n Eq. (13). The Poincaré section used here is the (n-1)-dimensional plane defined by fixing the first coordinate

$$y_i(1) = 0.$$

Let $\{\hat{y}_j\}_{j=1}^{\hat{N}}$ denote the collection of (n-1)-dimensional points on the Poincaré section obtained from the reconstructed trajectory $\{y_i\}_{i=1}^{N}$ with the additional condition that $y_i(1)$ is increasing, and define the integer function $i(j)$ by

$$P : y_{i(j)} \rightarrow \hat{y}_j \ ,$$

where P is the Poincaré projection. We define the expected return time by

$$\begin{aligned} \bar{i} &= \frac{1}{\hat{N}-1} \sum_{j=1}^{\hat{N}-1} (i(j+1) - i(j)) \\ &= \frac{1}{\hat{N}-1} (i(\hat{N}) - i(1)) \, . \end{aligned} \tag{33}$$

Like entropy, a Lyapunov exponent is a measure of predictibility, the major difference being that Lyapunov exponents are differential measures rather than integral. Next, we describe how a Lyapunov exponent is estimated. Let

$$\{(\hat{y}_{j_1(k)}, \hat{y}_{j_2(k)})\}_{k=1}^{\hat{K}}$$

be the collection of distinct pairs from the reconstructed Poincaré section such that $\hat{y}_{j_2(k)}$ is the closest point to $\hat{y}_{j_1(k)}$ on the section. Define

$$\bar{\lambda}_k = \frac{1}{\bar{i}} ln\{\frac{|\hat{y}_{j_2(k)+1} - \hat{y}_{j_1(k)+1}|}{|\hat{y}_{j_2(k)} - \hat{y}_{j_1(k)}|}\} \, ,$$

and the expected value

$$\bar{\lambda} = \frac{1}{\hat{K}} \sum_{k=1}^{\hat{K}} \bar{\lambda}_k \, . \tag{34}$$

For the example file used in this section, namely, $fe1ch13$ listed in Table 1, we obtain

$$\begin{aligned} \hat{K} &= 608 \, , \\ \bar{i} &= 72\, i \, , \\ \bar{\lambda} &= .22(-1)\,(e)/i \, . \end{aligned} \tag{35}$$

We note that as a result of conditioning the data Eq. (4), values for these parameters are not unduely affected by noisy, high frequency intersections with the Poincaré section. Referring to the entropy illustration Eq. (32), the time scale used for estimating entropy is $\bar{i}$ as defined by Eq. (33).

Note that the units of $\bar{\lambda}$ are the same as those of entropy, K. Both entropy and the Lyapunov exponents give a measure of exponential rate-of-loss of predictability. To

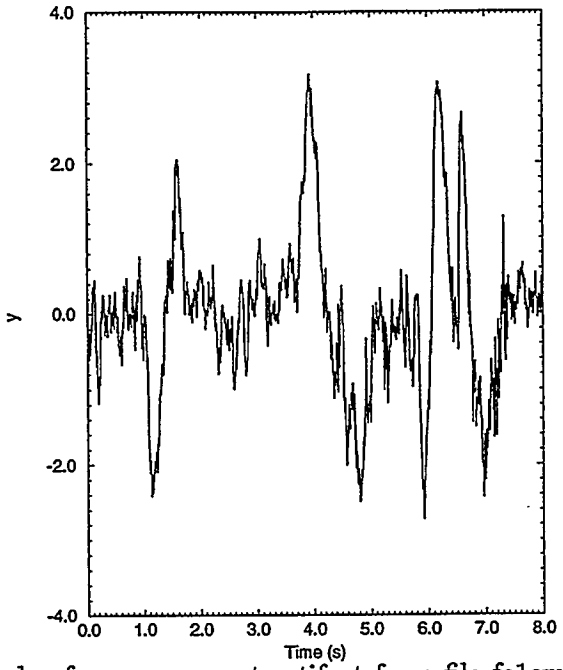

Figure 8: An example of eye-movement artifact from file fe1em13: EEG data, y, as a function of time.

repeat from earlier, if

$$|u_i - v_i| \propto e^{\lambda i},\ i = 0, 1, 2, \ldots,$$

where u_i, v_i are trajectories in state space, we say the trajectories are diverging at the exponential rate λ and refer to the units as units(λ)=(e)/i.

3. An Eye-Movement Artifact Filter

In this section we introduce a filter for removing eye-movement artifact from EEG time series. Eye movement causes low frequency activity in the time series data, as shown in Fig. 8. Referring to Table 1, this figure shows eight seconds of time series data with eye movement artifact from the first segment of file $fe1em13$.

The approach adopted here to remove low frequency artifact is to develop a low frequency, zero phase shift filter which follows the behavior of the artifact. This value is subtracted from the original signal and the result is an artifact free signal. The zero phase shift filter is generated using a least-squares estimation criterion to fit a quadratic polynomial to a moving time window of the time series data. Let the time series be denoted

$$\{y_i\}_{i=0}^{N} .$$

We define the window centered at index value j by

$$\{j - n + k\}_{k=0}^{2n} ,$$

where the parameter n is used to select a frequency band for the filter. Data in the window centered at index j, that is, $\{y_{j-n+k}\}_{k=0}^{2n}$, is approximated by the quadratic

expression

$$F_k = a_1(k-n)^2 + a_2(k-n) + a_3\ ,\ k = 0, 1, \ldots, 2n\ .$$

Define the quadratic functional

$$L(a_1, a_2, a_3) = \sum_{k=0}^{2n} [F_k - y_{j-n+k}]^2\ .$$

The optimal values for the coefficients $a_l, l = 1, 2, 3$ are obtained by minimizing L, and those values correspond to the solution of the linear syatem of equations

$$\frac{\partial L}{\partial a_l} = 0,\ l = 1, 2, 3\ .$$

The optimal polynomial evaluated at the center point $k = n$, which corresponds to the data value y_j, is denoted F_n^j. We denote the filter transformation by

$$F : y \rightarrow z\ , \tag{36}$$

where

$$z_j = y_j - F_n^j\ .$$

The sample time is $t_s = 1/f_s$, where $f_s = 512\,Hz$ is the sample frequency. Then, the time window parameter $n = 128$ determines the filter frequency

$$f_a = 1/(2n \times t_s)\ Hz = 2.0\ Hz\ , \tag{37}$$

which is tuned to the dominant frequency of eye-movement. Fig. 9 exhibits the result of applying the artifact filter to a segment of data containing eye-movement artifact. As for Fig. 8, the data we use is file $fe1emc13$ listed in Table 1. Fig. 9 displays power spectra of a segment of data with eye-movement and of the artifact filtered product of that segment. Also included in that figure is the power spectrum of an artifact-free segment from the file $fe1ch13$.

The frequency range for EEG signals presently considered to be neurophysiologically meaningfull is, approximately, $1Hz \leq f \leq 30Hz$, which is divided into the four energy bands

$$\begin{array}{lrcl} \delta: & & f < & 4\,Hz\ , \\ \theta: & 4\,Hz < & f < & 8\,Hz\ , \\ \alpha: & 8\,Hz < & f < & 13\,Hz\ , \\ \beta: & 13\,Hz < & f & \ . \end{array} \tag{38}$$

Eye-movement artifact falls in the δ-band. The main problem to be considered in assessing the effectiveness of the eye-movement artifact filter described here is that some of the energy in the δ-band that is not eye-movement related is also removed. However, referring to Fig. 9, the effect on signal strength in the θ, α, and β-bands is negligible.

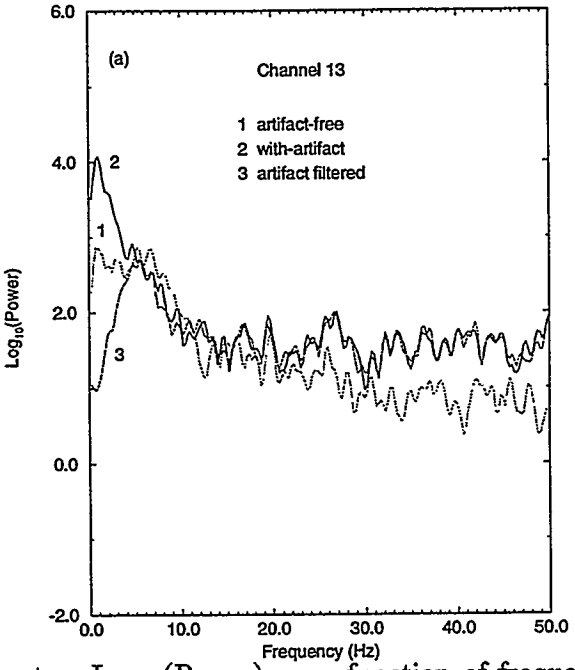

Figure 9: Power spectra, $\text{Log}_{10}(\text{Power})$, as a function of frequency for data with eye movement (y), eye movement filtered data (z), and the residual data (w).

4. Modeling and Analysis Results

In this section we present the results of the nonlinear analyses of EEG data. Three questions are considered. First, we look at the variability in results using artifact-free data observed by Channel 13. Second, we consider whether there is a difference between the dynamics observed at Channels 13 and 14 and whether the presence of artifact effects the results. As part of that study, we also look at the effect of the artifact filter presented in Section 3. Finally, we consider if there is coupling of information between Channels 13 and 14 and how that is effected by the presence of artifact and application of the artifact filter.

Before proceeding to address the three questions outlined above, we make a brief survey of the files listed in Table 1. Recall that each file is made up of several segments of time series data extracted from an EEG record. A data segment is denoted $\{y_i\}_{i=1}^{N}$, so that N is the number of observations in a segment. The minimum segment value is

$$y_{min} = \inf_{1 \le i \le N} \{y_i\} \; ; \tag{39}$$

the maximum segment value is

$$y_{max} = \sup_{1 \le i \le N} \{y_i\} \; ; \tag{40}$$

the mean value is

$$\mu_y = \sum_{i=0}^{N} y_i/(N+1) \; ; \tag{41}$$

the absolute average deviation is

$$\xi_y = \sum_{i=0}^{N} |y_i - \mu_y|/(N+1)\,; \tag{42}$$

and the standard deviation is

$$\sigma_y = [\sum_{i=0}^{N} (y_i - \mu_y)^2/N]^{1/2}\,. \tag{43}$$

Values for these parameters, including the length N, for each segment from the files listed in Table 1 are recorded in Table 2.

4.1. Variation in Non-Seizure, Artifact-Free Cases

We begin by examining non-seizure, artifact-free data observed by Channel 13 for Patients 1-4. Referring to Table 1, the data files we use are

$$\begin{aligned} File\,1 &= fe1ch13\,, \\ File\,2 &= 25402c13\,, \\ File\,3 &= 28049c13\,, \\ File\,4 &= 28158c13\,. \end{aligned} \tag{44}$$

File 1 is the one used in Section 2 for illustration.

Figure 10 displays power spectra for the four data sets. Note that for each case the as-measured data is conditioned using the low-pass filter Eq. (4), where the filter cutoff value $f_c = 50Hz$ is used for File 1 and the value $f_c = 60Hz$ is used for Files 2-4. Comparing power spectra for Files 1-4, note there is significantly more energy in the frequency range 20-30 Hz relative to the range 0-10Hz for Files 2,3 than for Files 1,4. The filter cutoff frequency $f_c = 60$ Hz preserves the energy content of Files 2,3 in a neighborhood of 20 Hz while at the same time isolating the band 1-30 Hz from frequencies greater than 30 Hz. Clearly, the dominant low frequency band is isolated in the conditioned data for Files 1,4. The reason for using the cutoff value $f_c = 60Hz$ rather than $f_c = 50Hz$ for File 4 is the rate of power loss with frequency over the high frequency face of the low frequency power band is not quite as sharp as that for File 1. For the remainder of this section we use the conditioned data.

Figure 11 shows results for the mutual information analyses of Files 1-4. The curves corresponding to Files 2,3 display more high frequency structure than those for Files 1,4, which is consistent with the observed power spectra structure of Files 2,3 in the neighborhood of 20 Hz compared to that of Files 1,4. Power spectra and mutual information appear to discriminate between Files 1-4 by grouping Files 1,4 and Files 2,3.

The power spectra and mutual information results determine consistent estimates for relevant time scales in each case. Based on those estimates, we choose the candidate parameter values $(n,k) = (9,3)$ for File 1 and $(n,k) = (8,3)$ for Files 2-4, which determine the window frequencies $f_w \approx 21Hz,\ 25Hz$ Eq. (15), respectively. Fig. 12

Dataset Name	Segment Index	N	y_{min}	y_{max}	μ_y	ξ_y	σ_y
fe1ch13	1	4096	-45	69	-0.6	11.3	14.0
	2	5120	-45	36	-0.2	9.5	11.9
	3	10496	-38	40	-0.7	8.3	10.5
	4	30208	-54	47	-0.6	9.0	11.4
fe1ch14	1	4096	-68	29	-2.7	6.6	8.4
	2	5120	-39	24	-2.8	6.2	7.9
	3	10496	-37	30	-2.7	6.0	7.7
	4	30208	-37	27	-2.8	6.1	7.8
25402c13	1	17920	-80	61	0.77	14.8	18.6
	2	9728	-69	60	0.71	13.6	17.0
	3	13824	-77	248	0.64	13.4	17.0
	4	12800	-45	38	0.78	8.6	10.9
28049c13	1	7680	-29	29	0.20	5.7	7.2
	2	5120	-26	26	0.14	5.4	6.8
	3	5120	-28	29	0.29	5.4	6.9
	4	6656	-26	37	0.20	6.1	7.8
	5	15360	-32	35	0.08	6.4	8.2
28158c13	1	10752	-15	53	14.6	6.6	8.3
	2	6656	-11	46	14.5	6.5	8.1
	3	7168	-23	109	14.8	7.5	10.7
	4	9728	-23	50	14.5	7.1	9.0
fe1emc13	1	3895	-42	40	0.12	8.7	11.1
	2	3895	-440	63	0.01	8.1	14.8
	3	3895	-45	44	0.08	8.1	10.6
	4	10039	-66	81	-0.02	7.4	9.9
fe1emc14	1	3895	-34	38	-0.05	5.6	7.5
	2	3895	-57	440	0.07	5.7	12.9
	3	3895	-31	24	-0.01	5.5	7.2
	4	10039	-72	44	0.00	6.0	8.0

Table 2: Segment statistics for files listed in Table 1.

displays the singular value analyses results. The distributions of singular values for Files 1-4 are not significantly different, although the slope of the distribution for File 1 is greater than the slopes for Files 2-4, which are on the whole very similar. We expect the difference to be revealed in the correlation integrals by a difference in state-space length scales where structure is resolved, as illustrated by Figs. 5,6.

Based on power spectra, mutual information, and singular values, and using the modeling technique described in Section 2.5, the candidate models we choose to represent Files 1-4 are

$$\begin{array}{lll} \textit{File}\ 1 & (n,k,m) = (9,3,9)\,, \\ \textit{File}\ 2,3,4 & (n,k,m) = (8,3,8)\,. \end{array} \tag{45}$$

From a modeling standpoint, the candidate models for Files 1-4 are very similar.

Correlation integrals for the four reconstructions are displayed in Fig. 13, and results for the estimated correlation dimensions are summarized in Table 3. First, note that low-dimensional information appears to be resolved by the candidate models. Second, note that the structure for Files 2-3 is over a range of state-space length scales translated in the positive direction relative to that for File 1, which is consistent with the distributions of singular values Fig. 12. The correlation dimension for File 1 is $\nu = 4.1$, which is significantly less than the range of values $4.8 \leq \nu \leq 5.2$ corresponding to Files 2-4.

Recall that we use the correlation integral to judge the completeness of a candidate model. Combining the correlation integral results with those for power spectra, mutual information, and singular values, we conclude the candidate models Eq. (45) are sufficiently complete to proceed with further analyses. While the correlation dimension of File 1 is distinguished from those values for Files 2-4, still the number of degrees-of-freedom in each process as represented by the models is approximately invariant, namely, five. We conclude that from a modeling standpoint, without additional information gained from analysing the models, Files 1-4 are very similar.

Recall that we use the time scale $\bar{i}$, defined by Eq. (33) as the average time between intersections of a trajectory with a Poincaré section, to analyze the entropy of a nonlinear dynamical process. For File 1, that expected return time is $\bar{i} \approx 72i$, whereas for Files 2,3,4 the expected return time varies as $23i$, $29i$, $38i$, respectively. Those values are summarized in Table 4. Note that the return times reflect the power spectra Fig. 10, the two longest periods corresponding to Files 1,4, where power is concentrated in the 0-10 Hz band, and the two shortest periods to Files 2,3, where there is significant power in a neighborhood of 20 Hz relative to that in the 0-10 Hz band.

Figure 14 displays the entropy dimension analysis results for Files 1-4, and Fig. 15 illustrates the corresponding characteristic wave speed calculations. A value for entropy is estimated from the values for entropy dimension and characteristic wave speed observed at time scale $\bar{i}$. The results are summarized in Table 3. The entropy values appear to separate the files into three ordered classes, where order corresponds to decreasing entropy. Those classes are

$$\begin{array}{ll} \textit{File}\ 2 & K \approx .66(-1)\,(e)/i\,, \\ \textit{File}\ 3,4 & K \approx .52(-1)\,(e)/i\,, \\ \textit{File}\ 1 & K \approx .31(-1)\,(e)/i\,. \end{array} \tag{46}$$

Table 4 summarizes results of the Poincaré return map analysis for the four cases. In each case, in addition to the expected return time, $\bar{i}$ Eq. (33), we estimate the Lyapunov exponent, $\bar{\lambda}$ Eq. (34). and the product $\bar{i} \times \bar{\lambda}$. Note the trend that, as $\bar{i}$ increases, $\bar{\lambda}$ decreases. Sequencing the files by decreasing value of $\bar{\lambda}$ produces an ordering consistent with that produced by entropy, but three classes are not distinguished as by entropy. The product $\bar{i} \times \bar{\lambda}$ does appear, however, to separate the files into the same three classes determined by entropy, but not with the same sense of order. The change in order appears to be due to the relatively large value of $\bar{i}$ for File 1. Ordering the classes according to decreasing value of that product results in

$$\begin{array}{ll} File\,1 & \bar{i} \times \bar{\lambda} \approx 1.58\,(e)\,, \\ File\,2 & \bar{i} \times \bar{\lambda} \approx 1.20\,(e)\,, \\ File\,3,4 & \bar{i} \times \bar{\lambda} \approx 1.15\,(e)\,. \end{array} \tag{47}$$

To judge overall variation, first consider the models that resulted for Files 1-4. Power spectra and mutual information, which provide consistent estimates of relavent time scales for each of the four cases, lead to the candidate modeling parameters Eq. (45). The singular value analysis results together with the correlation integral results imply the models based on those parameter values are complete. Also, the number of degrees-of-freedom revealed by the correlation integral analysis is approximately invariant across the four cases. From strictly a modeling standpoint, because the low-pass filter values used to condition the data and the parameters used to model the processes are approximately the same, we conclude the the processes represented by Files 1-4 are alike.

Next, consider the model analysis results. The correlation dimension for File 1 clearly distinguishes it from Files 2-4. Using Tables 3, 4, Fig. 16 displays the discrete values of entropy, K, the Lyapunov exponent, $\bar{\lambda}$, and the product $\bar{i} \times \bar{\lambda}$ plotted against the average return time, $\bar{i}$. Also included in that figure are three continuous curves. Consider the regression forms

$$y = a_0 + a_1 \bar{i} + a_2 \bar{i}^2\,, \tag{48}$$

$$y = \frac{a_0}{\bar{i}} + a_1 + a_2 \bar{i}\,. \tag{49}$$

The curve corresponding to $\bar{i} \times \bar{\lambda}$ is the result of modeling those values using the quadratic regression form Eq. (48). The curve corresponding to $\bar{\lambda}$ is the model Eq. (49) using the coefficients obtained from modeling $\bar{i} \times \bar{\lambda}$. Finally, the curve corresponding to K is the translation of the curve corresponding to $\bar{\lambda}$ determined by minimizing the mean-square error on K values. The coefficients resulting from the regression of $\bar{i} \times \bar{\lambda}$ on $\bar{i}$ using Eq. (48) are

$$a_0 = 1.593\,,\ a_1 = -0.2517 \times 10^{-1}\,,\ a_2 = 0.3472 \times 10^{-3}\,. \tag{50}$$

The coefficient of multiple determination [14] of that regression is

$$R^2 = 0.99974\,,$$

meaning that variability in the calculated values is virtually all accounted for. Note the relatively small value of the quadratic coefficient, a_2. Thus, the models reveal a variation in $\bar{\lambda}$ and K over the range of return time values $[23 \leq \bar{i} \leq 72]$ that is dominated by $\bar{i}^{-1}$.

	Patient			
	1	2	3	4
ν	4.1	5.2	4.8	5.2
s	4.5	6.0	5.5	5.0
ω	.67(-2)	.11(-1)	.93(-2)	.11(-1)
K	.31(-1)	.66(-1)	.52(-1)	.53(-1)

Table 3: Summary of correlation dimension and entropy analyses for normal, artifact-free data observed by Channel 13.

	Patient			
	1	2	3	4
$\bar{i}$	72	23	29	38
$\bar{\lambda}$	.22(-1)	.52(-1)	.39(-1)	.30(-1)
$\bar{i} \times \bar{\lambda}$	1.58	1.20	1.15	1.14

Table 4: Summary of Poincaré return-map analyses for normal, artifact-free observed by Channel 13.

4.2. The Dynamics Viewed from Distinct Channels

Next, we consider the process as observed from two distinct channels. The analyses presented here include consideration of the presence of eye-movement artifact and of the artifact filter. In the Introduction we described how two files are created from a given channel of the EEG record, one consisting of time segments free of all artifact and a second made up of time segments having only eye-movement artifact, where the segments in the latter file are strictly nonoverlapping in time with those in the former one. To analyse the eye-movement contaminated process, we create a third file by concatinating the first two. Finally, to examine the effect of the artifact filter we produce a fourth file by applying the filter F Eq. (36) to the third. In the present study, we use the eye-movement artifact files obtained from Channels 13 and 14 for Patient 1(see Table 1). We refer below to three data types, namely, artifact-free, with-artifact, and artifact-filtered. The three files that we analyse corresponding to Channel 13 can be summarized as

$$\begin{aligned} File\,1 &= fe1ch13\,, \\ File\,2 &= fe1ch13 + fe1em13\,, \\ File\,3 &= F(fe1ch13 + fe1em13)\,, \end{aligned} \tag{51}$$

with a similar description for Channel 14.

Figures 17,18,19 display power spectra, mutual information, and singular values for the six cases described above. Note that Fig. 17 includes power spectra for the

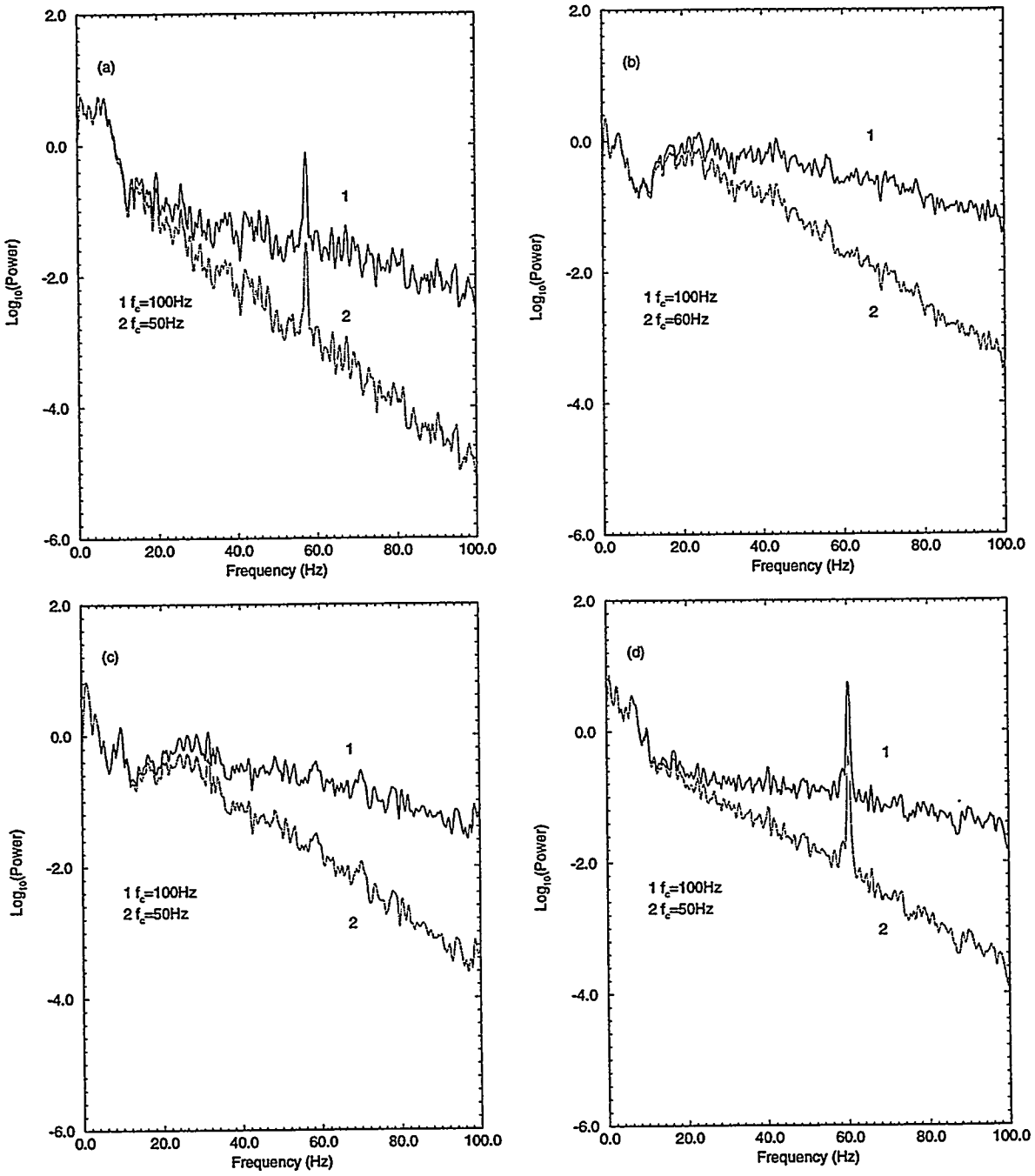

Figure 10: Power spectra, Log_{10}(Power), as a function of frequency for artifact-free cases: (a) Patient 1; (b) Patient 2; (c) Patient 3; (d) Patient 4.

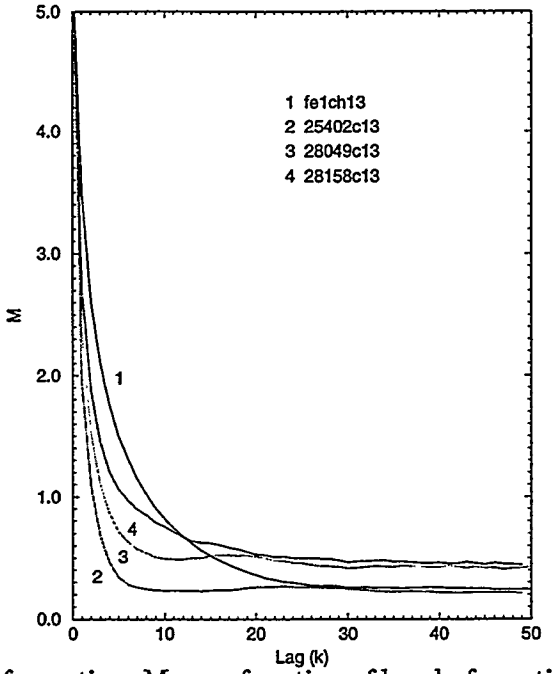

Figure 11: Mutual information, M, as a function of lag, k, for artifact-free data observed by Channel 13.

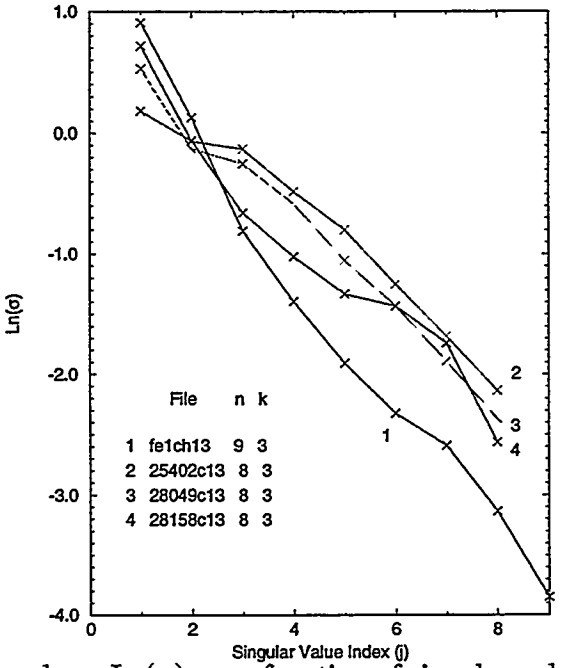

Figure 12: Singular values, Ln(σ), as a function of singular value index, j, for artifact-free data observed by Channel 13.

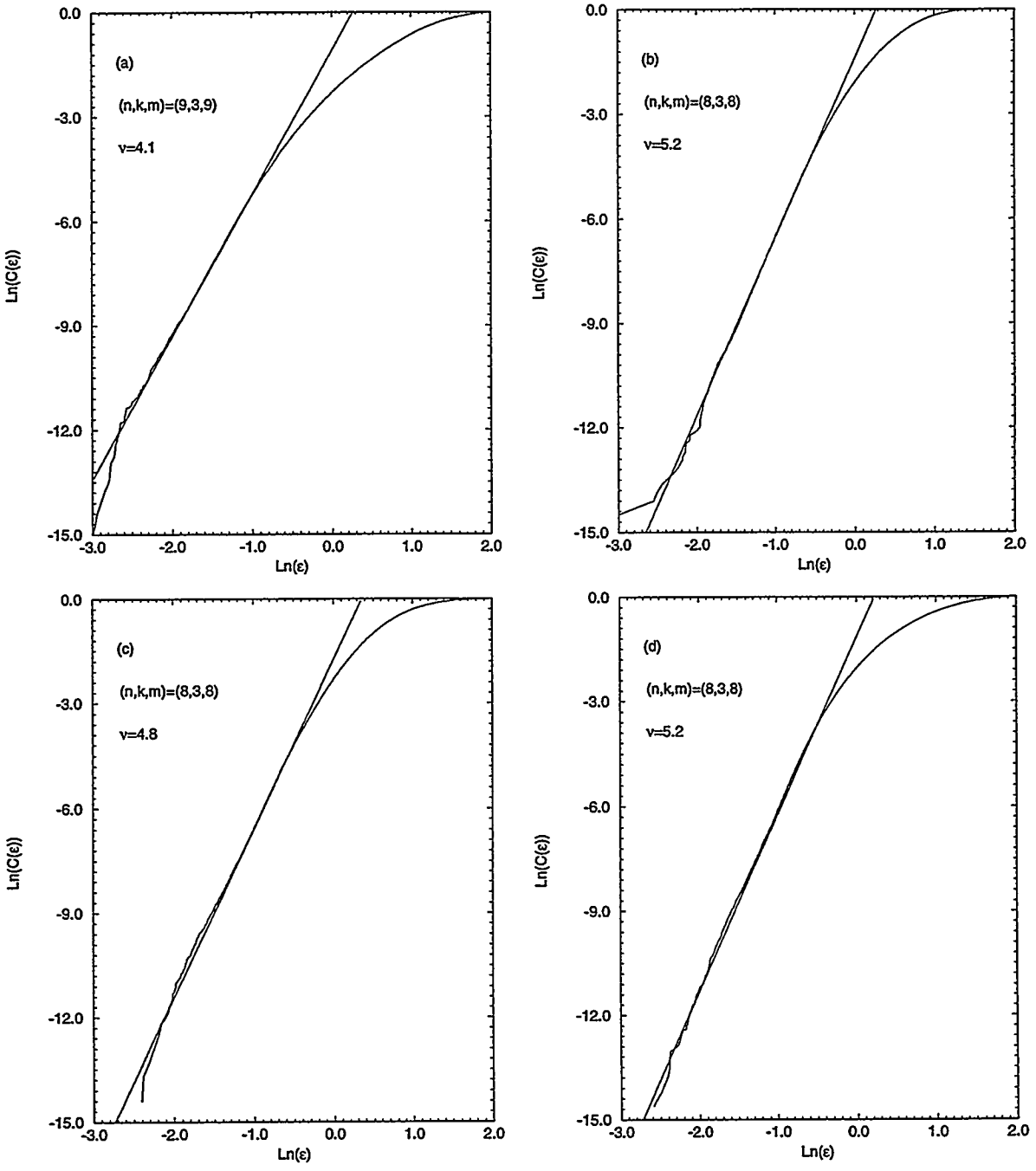

Figure 13: Correlation integrals, Ln(C(ϵ)), as a function of the partition, Ln(ϵ), for artifact free cases: (a) Patient 1; (b) Patient 2; (c) Patient 3; (d) Patient 4.

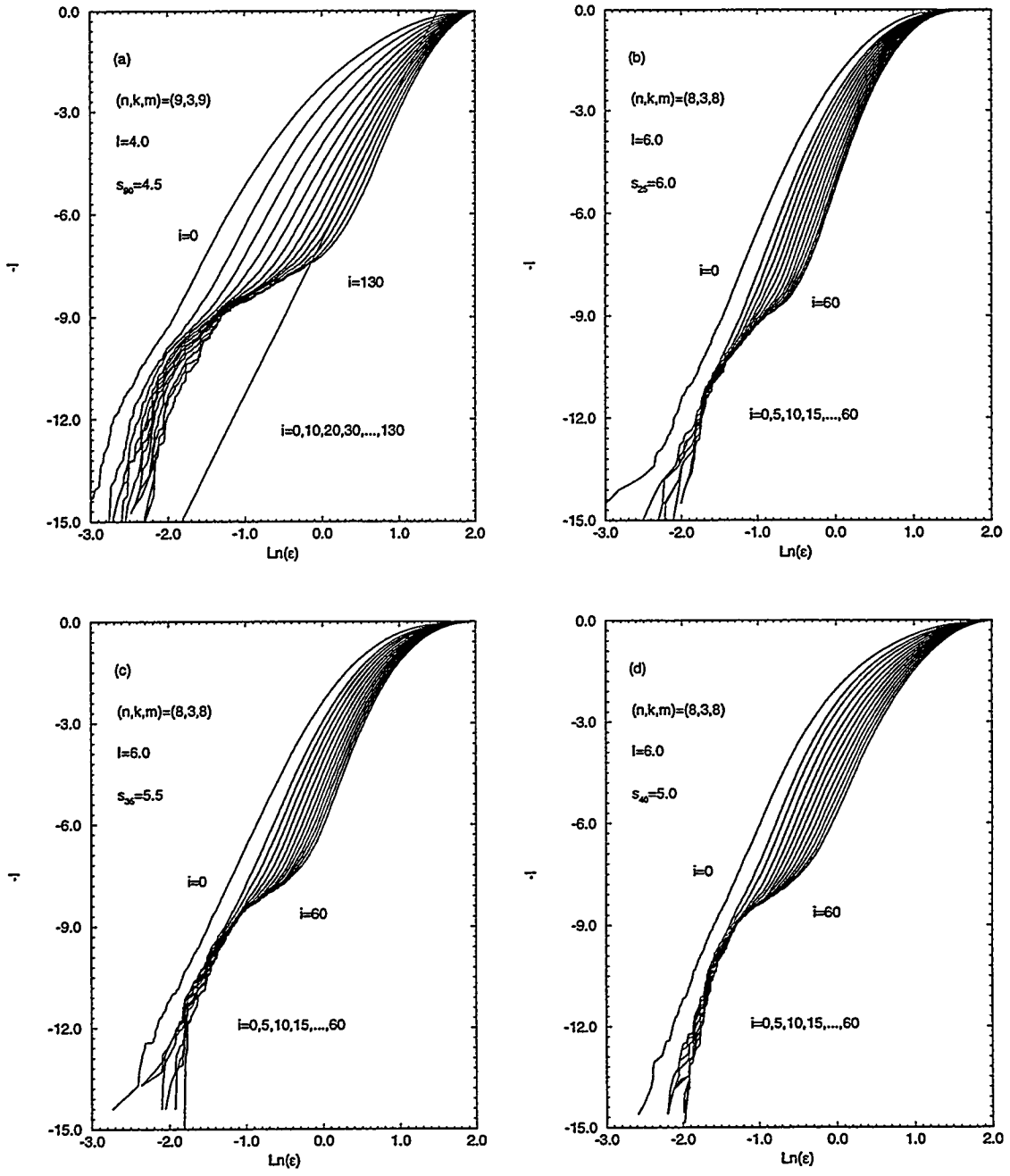

Figure 14: Estimation of entropy dimension, s, as the slope of $-\dot{I}(\mathrm{Ln}(\epsilon))$ for artifact free cases: (a) File 1(i=90); (b) File 2(i=25); (c) File 3(i=35); (d) File 4(i=40).

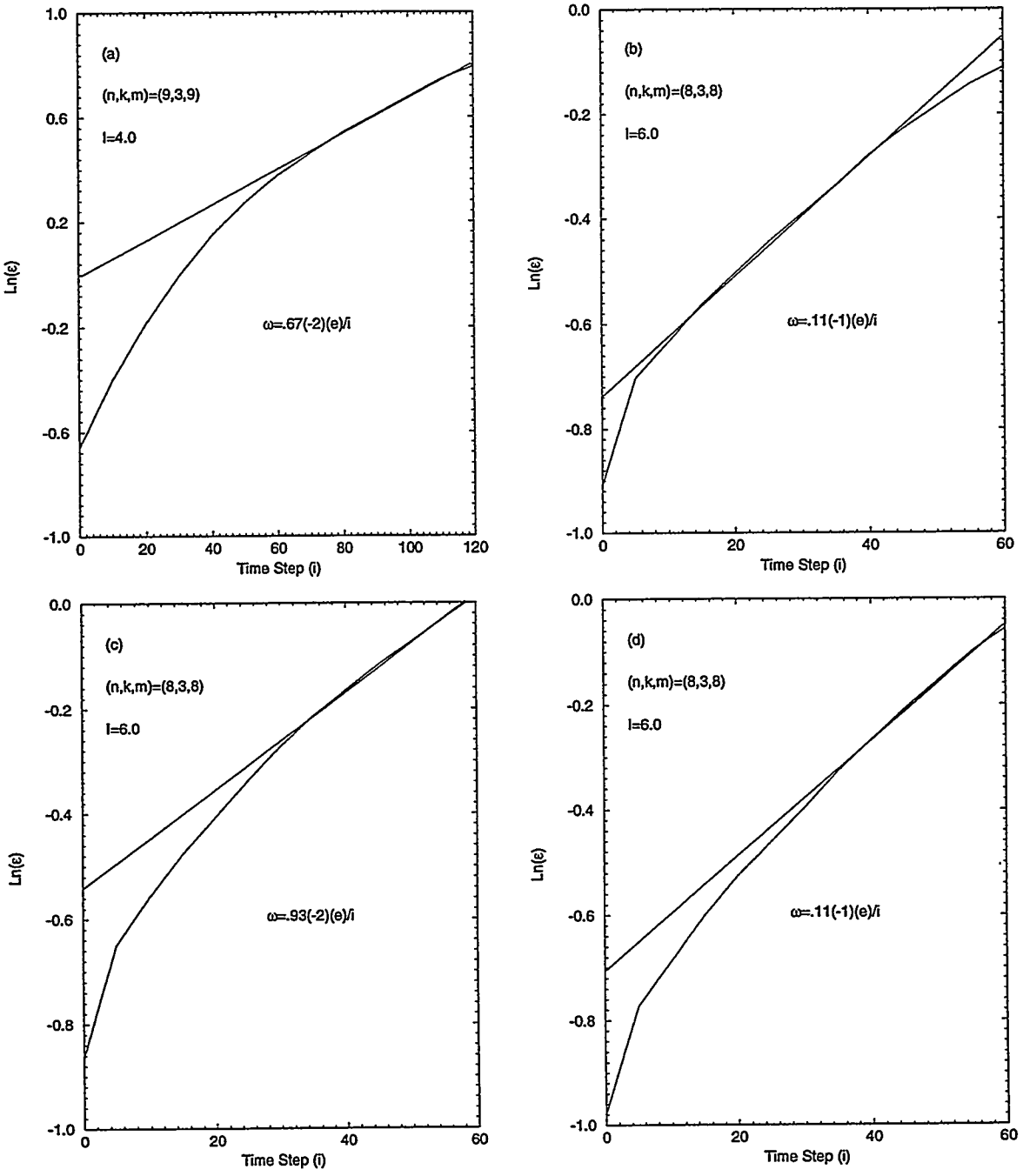

Figure 15: Estimation of characteristic wave speed, ω, as the constant rate of change of Ln(ϵ) with i using the relation $I = \dot{I}(\mathrm{Ln}(\epsilon))$ for artifact-free cases: (a) File 1(I=4.0); (b) File 2(I=6.0); (c) File 3(I=6.0); (d) File 4(I=6.0).

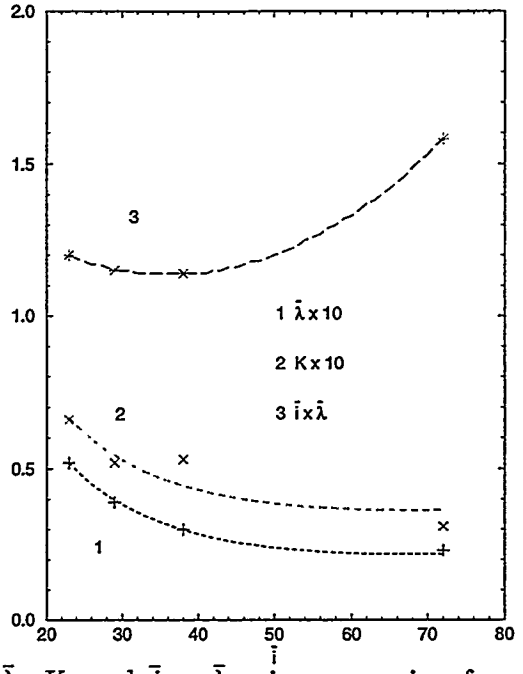

Figure 16: Modeling $\bar{\lambda}$, K, and $\bar{i} \times \bar{\lambda}$ using regression forms Eq. (49) and values displayed in Tables 3, 4.

conditioned datasets obtained using the low-pass filter cutoff value $f_c = 50$Hz, which is the same value used in the previous section for artifact-free data from Patient 1. The effect is, as before, to isolate the low frequency band of dominant power from the higher, noise-like frequencies. The mutual information results displayed in Fig. 18, as well as the singular values displayed in Fig. 19, are obtained using the conditioned data. For the remainder of this section, for each case we use the conditioned data.

Comparing results between Channels 13 and 14 for alike files, the power spectra, mutual information, and singular values are each very similar. Note that the addition of eye-movement results in a low frequency peak in the power spectra at approximately $f = 2$Hz that does not appear in the artifact-free power spectra. An effect resulting from the addition of eye-movement is also visible in the mutual information curves, but it is more pronounced in the Channel 13 case. That effect is to magnify the low frequency structure in mutual information, which is evidenced by the slightly lower rate of information loss with lag value. Conversely, the structure in mutual information for the artifact filtered data is more noise-like compared to the other two dataset types due to the reduced power in the low frequency band.

From the standpoint of constructing a representation of the dynamical processes, the power spectra, mutual information, and singular value analyses for each of the three data types, and for both channels, give results consistent with those found in Section 4.1 for the artifact-free data from Channel 13. Thus, the candidate modeling parameter values

$$(n, k, m) = (9, 3, 9) \tag{52}$$

used there Eq. (45) are used here for modeling all six cases.

Correlation integrals for the six cases are displayed in Fig. 20, and the resulting es-

timates for the correlation dimension are summarized in Table 5. First, the correlation dimension is resolved for all six cases, and therefore, we conclude the model Eq. (52) is adequate for each case. Referring to Table 5, the artifact-free, with-artifact, and artifact-filtered results are consistent both from the standpoint of an individual channel and between channels. For both channels, the correlation dimension estimates vary over the range of values ν =4.1-4.3. Note that the linear structure of the correlation integrals for the artifact filtered data, for both channels, is shifted toward larger values of $Ln(\epsilon)$ relative to the linear structure for the artifact-free and with-artifact cases. This is consistent with the singular value results displayed in Fig. 19.

Figures 21,22 illustrate the entropy dimension and characteristic wave speed results, respectively, for all six cases. The results are summarized in Table 5. Consider first Channel 13. The addition of eye-movement artifact increases the predictibility of the process, with entropy decreasing from $K = .31(-1)(e)/i$ to $K = .26(-1)(e)/i$. There is a smaller decrease in entropy going from the with-artifact case to the artifact-filtered case, where entropy for the artifact-filtered case is estimated to be $K = .24(-1)(e)/i$. Turning to the Channel 14 results, note that the estimates of entropy are relatively consistent across cases, the result being $K = .28(-1)(e)/i$ for both the artifact-free and with-artifact cases and $K = .26(-1)(e)/i$ for the artifact filtered case. The artifact-free process viewed from Channel 13 appears to be marginally more chaotic than that viewed from Channel 14. The presence of eye-movement artifact for Channel 13 has the effect of making the process appear marginally more stable, but not for Channel 14. There is a relative decrease in entropy of $7 - 8\,\%$ going from the with-artifact to the artifact filtered case for both channels.

The Poincaré return map estimates for the expected return time, $\bar{i}$, and Lyapunov exponent, $\bar{\lambda}$, for artifact-free, with-artifact, and artifact-filtered data are summarized in Table 6. The product $\bar{i} \times \bar{\lambda}$ is also shown in that table. The Poincaré return map estimates for each channel are consistent for artifact-free versus with-artifact data, but significantly different in comparison to artifact-filtered data. For each channel, going from artifact-free, or with-artifact, to artifact-filtered data there is a decrease in $\bar{i}$ of approximately 50 % and a decrease in the product $\bar{i} \times \bar{\lambda}$ of approximately 25 %. Comparing channels, for artifact-free and with-artifact data there is a significant decrease in $\bar{i}$ going from Channel 13 to Channel 14, but the product $\bar{i} \times \bar{\lambda}$, is virtually invariant. The same relationships hold for the artifact filtered data going from Channel 13 to Channel 14.

Consider again power spectra and mutual information. Power spectra for corresponding data types are very similar across channels, but there are significant differences across types. The addition of artifact adds power to the low frequency peak of the artifact-free power spectra, and the artifact filter removes all the power above the noise floor in that peak. Concerning mutual information, the addition of artifact is marginally more pronounced for Channel 13 than for Channel 14, the effect being to decrease the rate of decay of information. The artifact filter causes an increase in that rate of decay, the effect again being more pronounced for Channel 13 than for Channel 14. From a linear point-of-view, the affect on mutual information of adding artifact implies an enhancment of predictability, while the result of applying the artifact filter implies a decrease in predictabilty.

We continue our consideration of the predictability results. The addition of artifact to artifact-free data for Channel 13 has the effect of maginally decreasing entropy, K, which is consistent with the linear interpretation of the effect on mutual information, namely, an enhancement in predictabilty. A change in entropy with the addition of artifact is not observed at Channel 14. On the other hand, the effect of the artifact filter on entropy, which is to decrease the value, is opposite to the linear interpretation of the effect of the artifact filter on mutual information.

The effect on predictabilty caused by the addition of artifact implies, for the single case studied here, that eye-movement artifact has a marginally linear impact on the process model. Aside from implications about effectiveness of the artifact filter vis-à-vis clinical diagnosis of EEG data, the artifact filtered case reveals something about the structure of the process being observed. First, removing frequency content in a band, in this case the δ-band, $f < 4.0Hz$ Eq. (38), does not affect the dimension. This is consistent with an observation by Lawkins et al. [13] that dimensional information is not necessarily lost by low-pass filtering, that the effect is only on the range of state-space length scales where structure is resolved, even though frequencies judged to be relevant to the process being observed may be affected. In turn, that is consistent with the view that the process is nonlinear. Further, consider the affect on predictability. Both entropy and the product of return time with the estimated Lyapunov exponent decrease, meaning predictabilty increases, as a result of removing low frequency content. This is consistent with the interpretation that the process is a chaotic nonlinear dynamical system.

Patient 1						
	Channel 13			Channel 14		
	Artifact Free	With Artifact	Artifact Filtered	Artifact Free	With Artifact	Artifact Filtered
ν	4.1	4.3	4.3	4.2	4.3	4.3
s	4.5	4.7	4.8	4.8	5.2	4.8
ω	.67(-2)	.55(-2)	.50(-2)	.58(-2)	.55(-2)	.46(-2)
K	.31(-1)	.26(-1)	.24(-1)	.28(-1)	.28(-1)	.26(-1)

Table 5: Correlation dimension and entropy results using artifact free, eye movement artifact, and artifact filtered data observed by Channels 13 and 14

4.3. Coupling Between Channels

Next, we consider the question of coupling between channels, which is to ask to what degree distinct channels are observing the same brain dynamics. To some extent we can address that question using the results concerning artifact found in Section 4.2 by simply comparing analyses for the two channels. A more objective and quantitative approach is to use mutual information. For this purpose, three pairs of files are used,

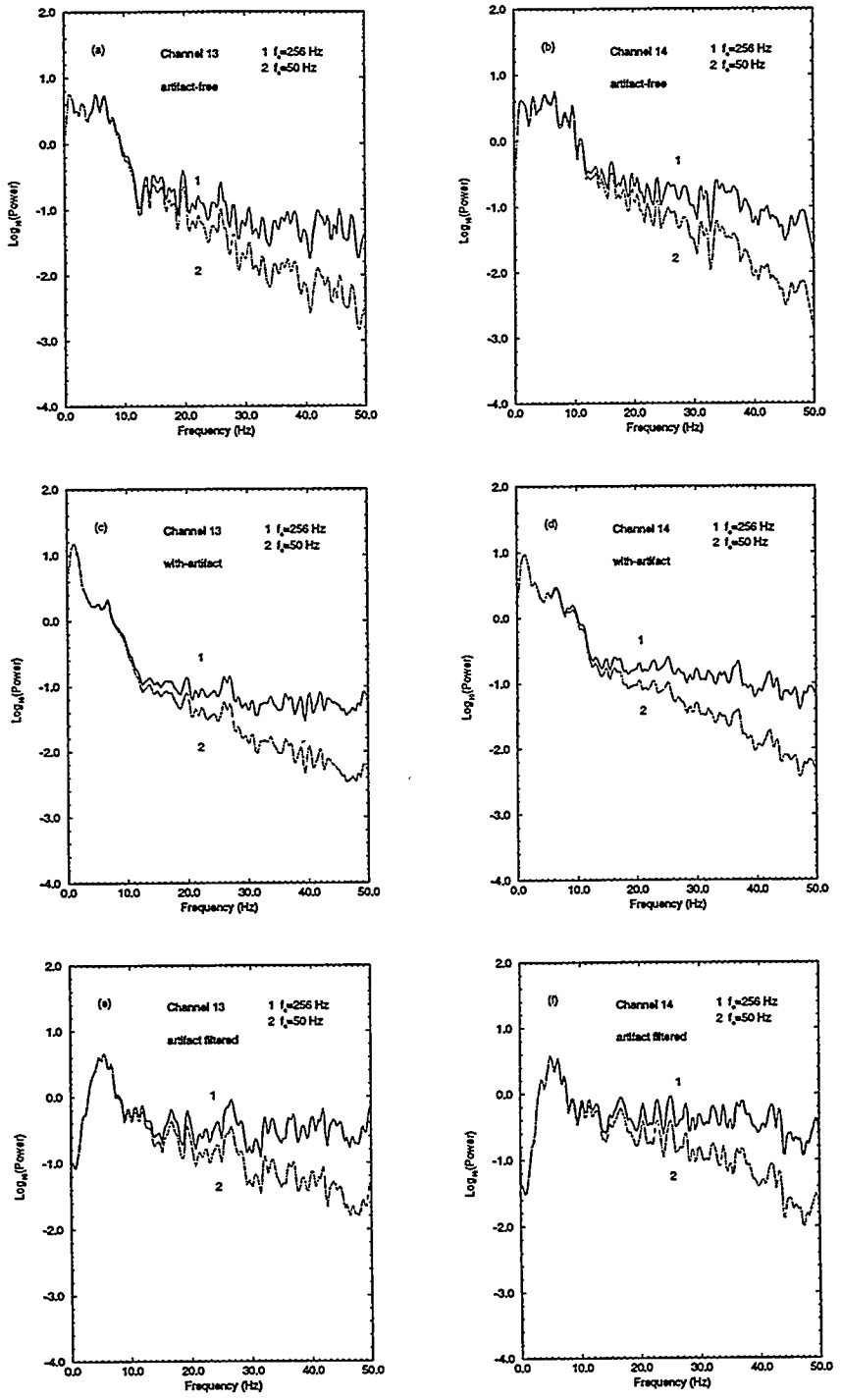

Figure 17: Power spectra, $\text{Log}_{10}(\text{Power})$, as a function of frequency for artifact-free, eye-movement artifact, and artifact filtered data: (a) artifact-free, channel 13; (b) artifact-free, channel 14; (c) eye-movement artifact, channel 13; (d) eye-movement artifact, channel 14; (e) artifact filtered, channel 13; (f) artifact filtered, channel 14.

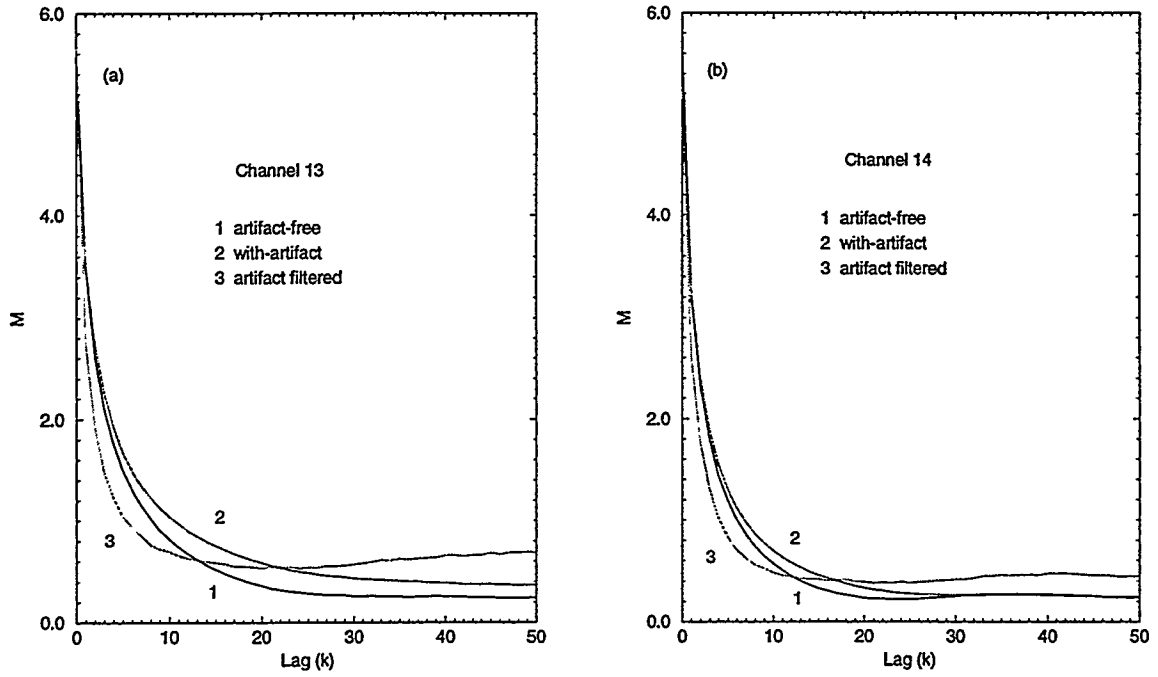

Figure 18: Mutual information, M, as a function of lag, k, for artifact-free, eye-movement artifact, and artifact filtered data: (a) channel 13; (b) channel 14.

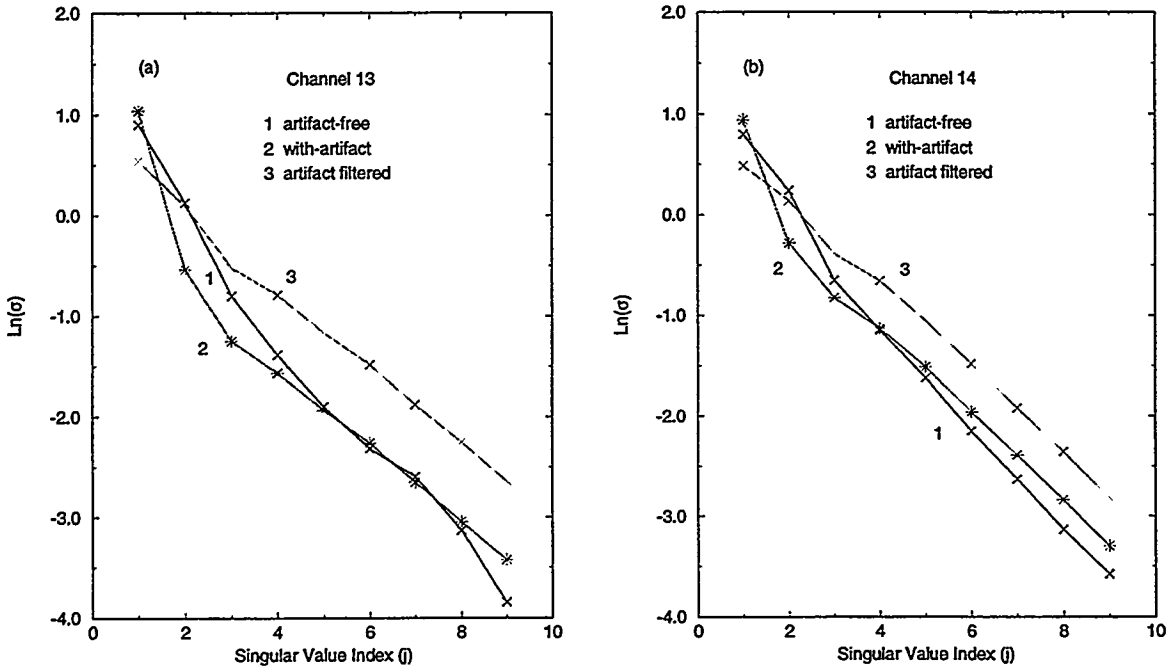

Figure 19: Singular values, Ln(σ), as a function of singular value index, j, for artifact-free, eye-movement artifact, and artifact filtered data: (a) channel 13; (b) channel 14.

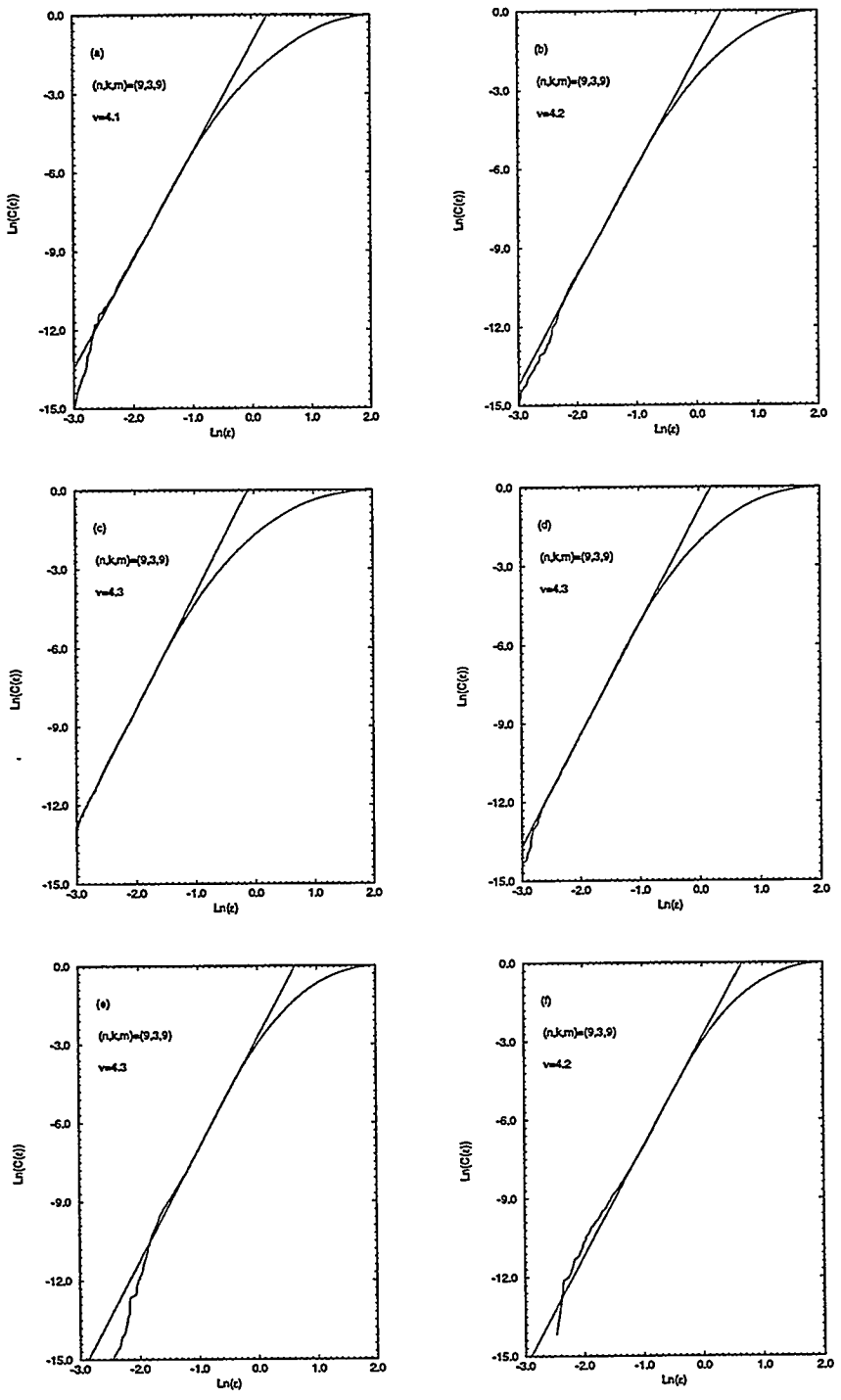

Figure 20: Correlation integrals, Ln(C(ϵ)), as a function of the partition, Ln(ϵ), for artifact free, eye movement artifact, and artifact filtered data: (a) artifact free, Channel 13; (b) artifact free, Channel 14; (c) eye movement artifact, Channel 13; (d) eye movement artifact, Channel 14; (e) artifact filtered, Channel 13; (f) artifact filtered, Channel 14.

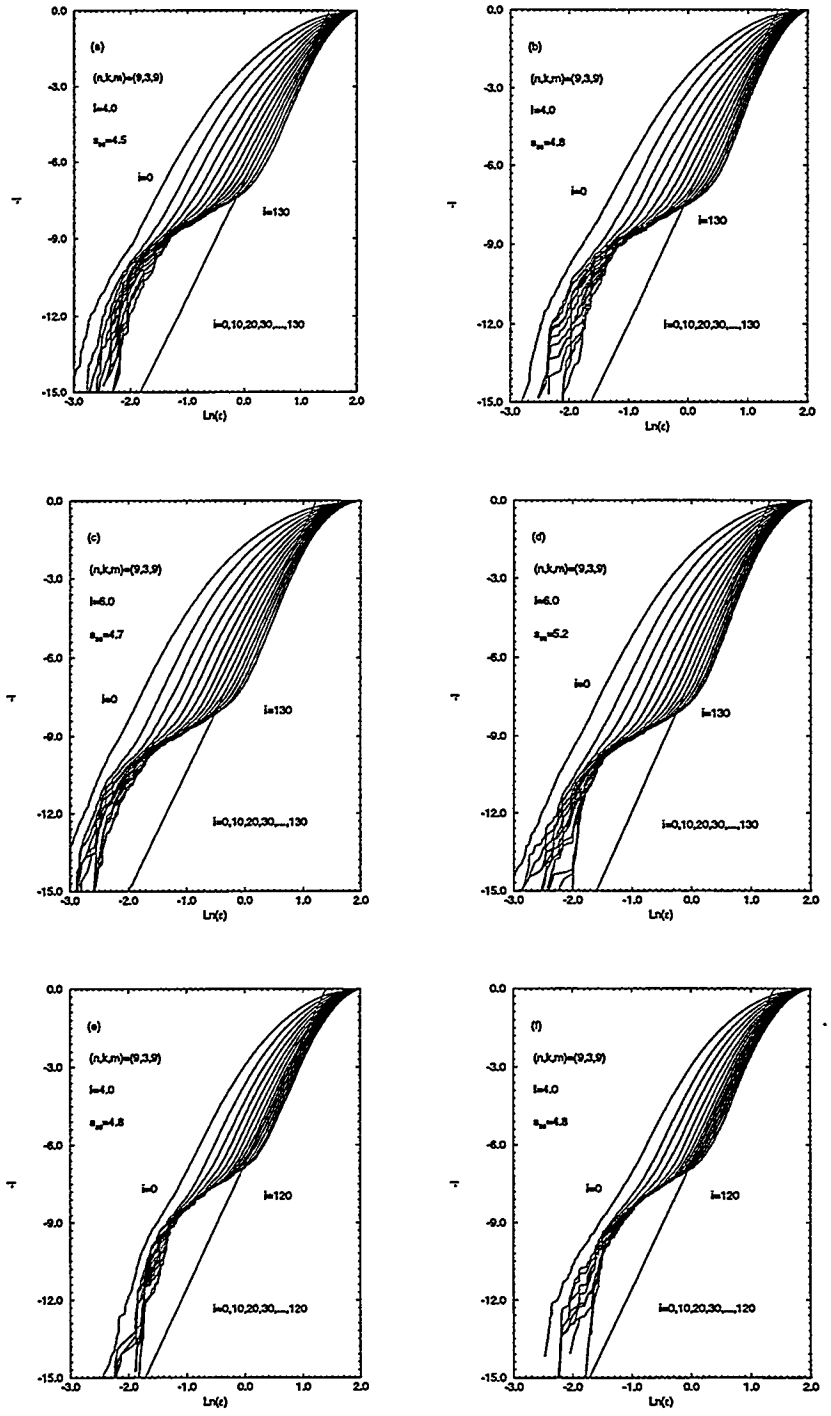

Figure 21: Estimation of entropy dimension, s, as the slope of $-\dot{I}(\mathrm{Ln}(\epsilon))$ for artifact free, eye movement artifact, and artifact filtered data: (a) artifact-free, Channel 13(i=90); (b) artifact-free, Channel 14(i=80); (c) eye-movement artifact, Channel 13(i=90); (d) eye-movement artifact, Channel 14(i=90); (e) artifact filtered, Channel 13(i=80); (f) artifact filtered, Channel 14(i=80).

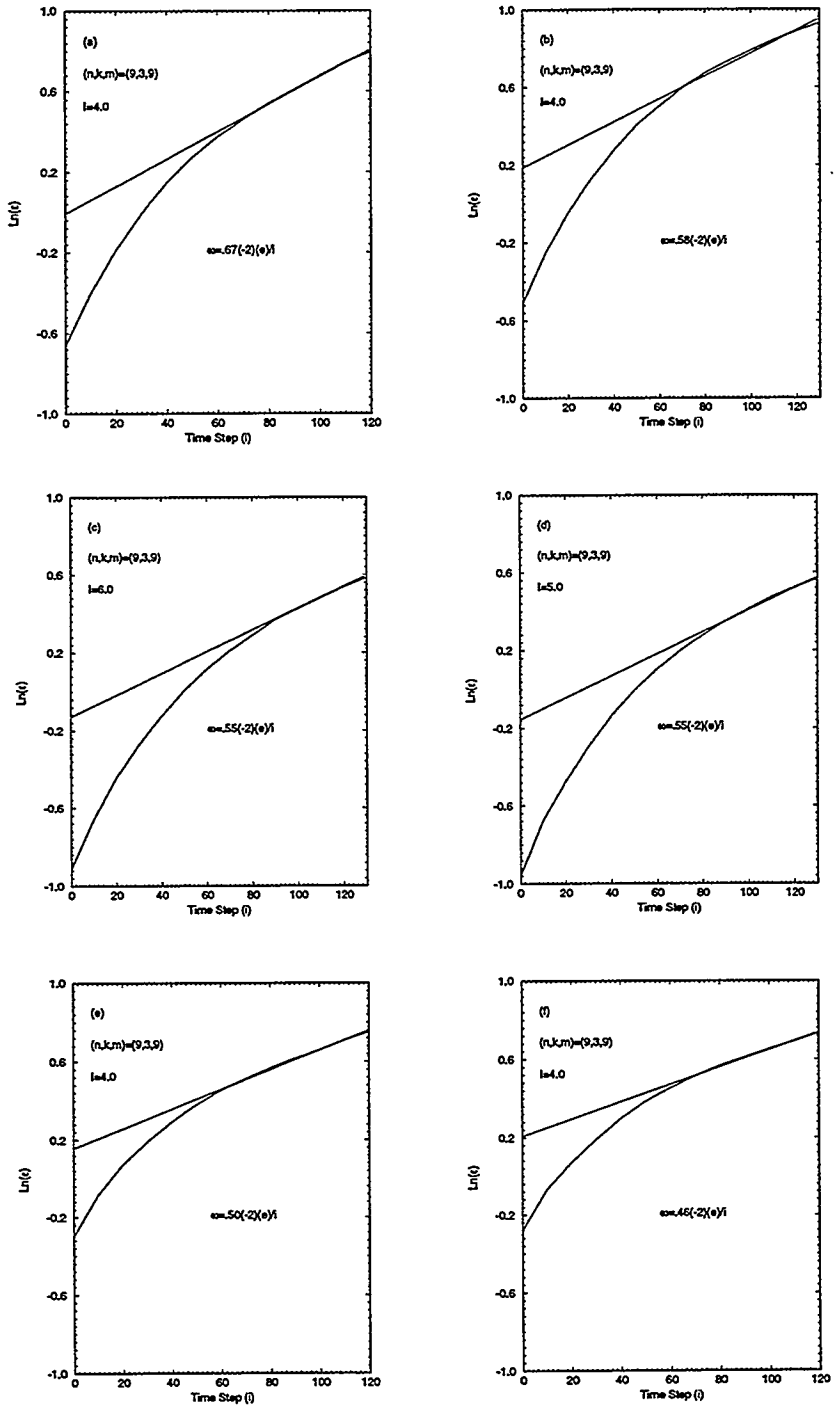

Figure 22: Estimation of characteristic wave speed, ω, as the constant rate of change of $\mathrm{Ln}(\epsilon)$ with i using the relation $\mathrm{I} = \dot{\mathrm{I}}(\mathrm{Ln}(\epsilon))$ for artifact-free, eye-movement artifact, and artifact filtered data: (a) artifact-free, Channel 13(I=4.0); (b) artifact-free, Channel 14(I=4.0); (c) eye-movement artifact, Channel 13(I=6.0); (d) eye-movement artifact, Channel 14(I=6.0); (e) artifact filtered, Channel 13(I=4.0); (f) artifact filtered, Channel 14(I=4.0).

Patient 1						
	Channel 13			Channel 14		
	Artifact Free	With Artifact	Artifact Filtered	Artifact Free	With Artifact	Artifact Filtered
$\bar{i}$	72	72	37	58	60	31
$\bar{\lambda}$	.22(-1)	.21(-1)	.32(-1)	.28(-1)	.26(-1)	.39(-1)
$\bar{i} \times \bar{\lambda}$	1.6	1.5	1.18	1.6	1.6	1.21

Table 6: Poincaré return map results using artifact-free, eye-movement artifact, and artifact filtered data observed by Channels 13 and 14

which are described as

$$\begin{aligned} File\,1 &= (fe1ch13, fe1ch14) \\ File\,2 &= (fe1em13, fe1em14) \\ File\,3 &= (F(fe1em13), F(fe1em14)), \end{aligned} \tag{53}$$

where F denotes the artifact filter Eq. (36). Referring to Table 1, recall that the file pairs File 1,2 consist of several distinct segments of data, but that those segments are alined in time.

Let x denote data from Channel 13 and y denote data from Channel 14. Referring to the definition of mutual information Eq. (11), for a fixed value of lag k, we estimate

$$M(k) = M(x_i, y_{i+k}) \tag{54}$$

and refer to it as "bivariate" mutual information. This statistic is estimated for both positive and negative time lag values. A relative maximum value of M for lag value $k = 0$ implies a source of activity is being observed simultaneously by Channels 13 and 14. A relative maximum for a negative value of lag implies a signal observed by Channel 14 is observed at Channel 13 after a delay corresponding to the absolute value of that lag, whereas a relative maximum for a positive lag value implies an observation at Channel 13 is related to an observation at Channel 14 at a later time corresponding to that positive lag value. The difference between the positive lag value for a relative maximum and the negative lag value of a relative maximum provides another a time scale for which we expect to see correlation between observations at each of the two channels.

The meaning of "coupling between channels" needs to be clarified. In order to heuristically describe the bivariate mutual information results presented here, we can think in terms of relatively large length-scale neurophysiological activity and relatively small length-scale activity. The relatively large length-scale activity is something we expect to observe simultaneously at two distinct channels, whereas small length-scale activity is local and is something we expect to observe directly only by a channel positioned at that location. However, local activity can be transmitted neurophysiologically across the extent of the brain, and consequently, with a time delay we may expect to observe the occurrence of local activity proximate to the location of one channel at another channel. There is also the possibility of a non-neurophysiological mechanism for

transmission, which is relavent here because of eye-movement artifact. While there is certainly neorophysiological activity associated with eye-movement, there is also muscular activity that is mechanicaaly transmitted through the skin covering the scalp. Recall from Fig. 1 that Channel 13 is located adjacent to the right eye and that Channel 14 is located one position further away along the hemispheric boundary.

Figure 23a shows the bivariate mutual information function between Channels 13 and 14 for artifact-free data. There are clearly defined relative maxima at lag values $k = 0$ and $k \approx -38$, and another relative maximum, albeit weakly defined, at $k \approx 40$. The band centered at $k = 0$, which is relatively symmetric, extends approximately over the range of lag values $-16 \leq k \leq 20$. While the relative maxima of the side bands are symmetrically located, the structures of those bands are significantly different, suggesting a greater sense of signal transmission from Channel 14 to 13 than from Channel 13 to 14.

Bivariate mutual information for the with-artifact data type is displayed in Fig. 23b. The central band for the with-artifact case is much broader than that for the artifact-free case, almost covering the range of lag values corresponding to both side-bands for the artifact-free case, and no side-bands are visible. Also, the magnitude of bivariate mutual information in this case is an order-of-magnitude greater than that for the artifact-free case. Further, note the difference in decay rate of mutual information for positive and negative lag values. This difference in decay rate suggests there is a difference in the dominant mode of transmission of eye-movement information between the two channels.

The bivariate mutual information for the eye-movement artifact filtered data is shown in Fig. 23c. In addition to the band centered at lag value $k = 0$, there are also side-bands for both positive and negative lag values. The central band displays a degree of nonsymmetry, the variation for negative lag values being greater than that for positive lags. The central band extends approximately over the range of lag values $-16 \leq k \leq 30$. The left side-band is split into two peaks, the relative maxima occuring at approximately the lag values $k = -36, -50$, and the relative maximum of the right side-band is approximately at the lag value $k = 52$.

We want to focus on comparing the artifact-free result with that for the artifact filtered case. Consider the central bands for those cases. The range of lag values for the two central bands, together with the frequencies associated with those values, are

$$\begin{array}{lll} \text{artifact-free:} & -16 \leq k \leq 20 \;\;,\;\; & 14.2\,Hz \leq f\,, \\ \text{artifact filtered:} & -16 \leq k \leq 30 \;\;,\;\; & 11.1\,Hz \leq f\,. \end{array} \tag{55}$$

Now, consider the side-bands. We estimate the range of lag values for the side-bands to be

$$\begin{array}{lll} \text{artifact-free:} & -60 \leq k \leq -16 \;\;,\;\; & 28 \leq k \leq 52\,, \\ \text{artifact filtered:} & -62 \leq k \leq -26 \;\;,\;\; & 34 \leq k \leq 58\,. \end{array} \tag{56}$$

The range of lag values between side-bands for the two cases, and the frequencies corresponding to those lag values, are

$$\begin{array}{lll} \text{artifact-free:} & 44 \leq k \leq 112 \;\;,\;\; & 4.6\,Hz \leq f \leq 11.6\,Hz\,, \\ \text{artifact filtered:} & 60 \leq k \leq 120 \;\;,\;\; & 4.3\,Hz \leq f \leq 8.5\,Hz\,. \end{array} \tag{57}$$

If the processes being observed were linear and linearly correlated, than we would expect a time scale of transmission to be associated with the frequency in the power spectra corresponding to that time scale. If, on the other hand, the processes are nonlinear, suggesting noncoherence of a frequency component oberved at either channel, or nonlinearly correlated, meaning there is a nonlinear relationship between time scales in the processes and time scales of transmission, then we would not expect to be able to compare a time scale of transmission with the corresponding frequency component of the process. We shall assume that the nonlinear structures are are not so strong to completely eliminate the underlying linear structures, so that we can learn something by comparing time scales from bivariate mutual information to time scales in the power spectra of the two channels.

We suppose that the central band corresponds to signals that simultaneously reach the two channel locations and that the band width is a time scale describing how long that signal is observed. Further, we suppose the side-bands represent time scales associated with the transmission of information from one channel to the other and back. Consider the power spectra for artifact-free and artifact filtered data from Channels 13 and 14 displayed in Fig. 17. Note in both cases the upper bound of the low frequency band in the power spectra isolated by the low-pass filter corresponds to the low frequency bound of the central band in bivariate mutual informatiion Eq. (55). It follows, given our supposition concerning the central band, that the time scales over which a signal simulataneously reaching the two channels continue to effect both channels correspond to "noise" in that signal, that is, to frequencies higher than the low frequency band of significant power in the power spectra. Further, note that for both cases the frequency range associated with the side-band Eq. (57) corresponds well with the primary power band displayed in the power spectrum, the only exception being for the lowest frequencies in the artifact-free case.

The artifact filtered data differs from the artifact-free data in two ways. First, the artifact filtered data inculdes some artifact information, and second, as it results from processing by the artifact filter, the artifact filtered data includes an effect due to that filter. For example, the central band of the artifact filtered case Fig. 23c displays a residual effect due to artifact resulting from the nonsymmetry in that band for the with-artifact data Fig. 23b. The artifact filter is a linear process, and we expect the artifact filtered data to display linear structure not evidenced in the artifact-free data. Consider the side-bands in the artifact-free Fig. 23a and artifact filtered cases. The pair of symmetric peaks for lag values $k \approx \pm 50$ in the artifact filtered case, which do not appear in the artifact-free case, is an example of the filter effect.

5. Conclusions

Consider the results from Section 4.1 for modeling and analyzing artifact-free data for Patients 1-4. Again, distinguishing between a model and the analysis results obtained using that model, a significant result is that there is very little variation in the models over the four cases. Also, it is significant that in each case we found low dimensional structure and that the number of degrees-of-freedom in each process accounting for that structure is approximately invariant. This result has the important implication

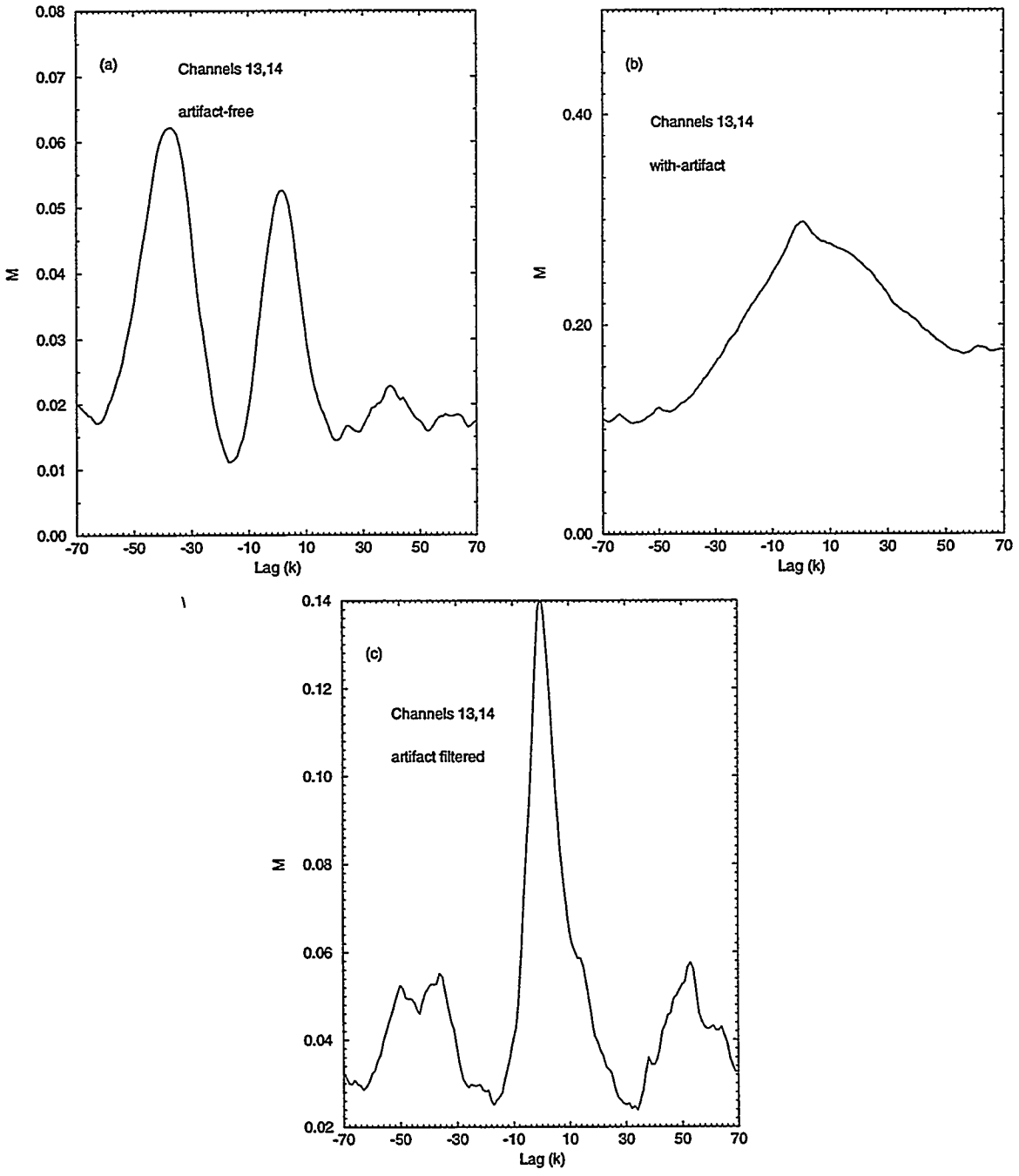

Figure 23: Bivariate mutual information, M Eq. (54), as a function of lag, k, for Channels 13,14 data: (a) artifact-free; (b) eye-movement artifact; (c) artifact filtered.

that only a few models may be needed to represent a relatively large number of cases. Concerning the analysis results obtained from the models, a significant finding is that nonlinear predictabilty parameters appear to regress with the average Poincaré return time and that over the range of return times represented by the four cases that regression is dominated by inverse return-time structure.

Consider now the results from Section 4.2 concerning artifact-free, with-artifact, and artifact filtered data from Channels 13 and 14 for Patient 1. First, we found that the models, and the degrees-of-freedom of each process resolved by those models, are approximately invariant over all six cases. Second, we found that for like data types, the differences between Channel 13 and 14 model analysis results are relatively small, the only major exception being the Poincaré return time. We focus on changes that occur in going from artifact-free data to with-artifact data and changes that occur in going from artifact-free data to artifact filtered data. Consider going from artifact-free to with-artifact data. In Channel 13 data we found that predictability increases with the addtion of artifact and that the effect of artifact appears to be linear. Further, we found there is virtually no change in predictability in the Channel 14 data. This result for the two channels is significant because it suggests that eye-movement artifact has a weak, linear effect on the models. In going from the artifact-free data to the artifact filtered data predictability increases significantly for both channels. The primary effect of eye-movement in frequency space is the low frequency δ-band, and the artifact filter severly reduces power in that band relative to the artifact-free data. This result implies the δ-band contributes significant nonlinear structure to the neurophysiological processes observed at Channels 13 and 14.

Consider finally the results from Section 4.3 concerning bivariate mutual information between Channels 13 and 14 for Patient 1 for the three data types. Concerning the artifact-free case, a significant result is that, while the level of mutual information is small, there is a quantitatve difference in transmission from Channel 14 to 13 compared to transmission from Channel 13 to 14. With the addition of artifact and application of the artifact filter, that difference appears to be lost, suggesting that Channel 13 and 14 dynamics may appear more alike for the artifact filtered case than for the artifact-free case. The addition of artifact alone to the artifact-free data has a strong linear effect, so strong that the bivariate mutual information structure observed in the artifact-free case is completely masked. Again, the results suggest that the dynamics for with-artifact data may appear more alike at the two channels than for the artifact-free case. Another significant result is that the time scales determined from the bivariate mutual information curves for the artifact-free and artifact filtered cases correlate well with the corresponding frequencies in the power spectra for like data types. This result implies that while the processes at the two channels may be nonlinear and transmission between channels may also be nonlinear, those nonlinearities are not so strong as to eliminate a linear relationship between the local time scales represented by power spectra and transmission time scales represented by bivariate mutual information.

6. References

[1] H.D.I. Abarbanel, R. Brown, J.J. Sidorowich, and L.S. Tsimring. The analysis of observed chaotic data in physical systems. *Rev. Mod. Phys.*, 65:1331–1392, 1993.

[2] A. Babloyantz and A. Destexhe. Low dimensional chaos in an instance of epilepsy. *Proc. Natl. Acad. Sci. USA*, 83:3513–3517, 1986.

[3] E. Basar(ed.). *Chaos in Brain Function.* Springer-Verlag, Berlin, 1990.

[4] A. Ben-Mizrachi, I. Procaccia, and P. Grassberger. Characterization of experimental (noisy) strange attractors. *Physical Review A*, 29:975–977, 1984.

[5] R. B. Blackman and J. W. Tukey. *The Measurement of Power Spectra.* Dover Publications, Inc., New York, NY, 1958.

[6] D. S. Broomhead and G. P. King. Extracting qualitative dynamics from experimental data. *Physica D*, 20:271, 1986.

[7] J.P. Eckmann and D. Ruelle. Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.*, 57:617, 1985.

[8] A.M. Fraser and H.L. Swinney. Independent coordinates for strange attractors from mutual information. *Physical Review A*, 33:1134–1140, 1986.

[9] M. Fraser. Reconstructing attractors from scalar time series: A comparison of singular systems and redundancy criteria. *Physica D*, 34:391–404, 1989.

[10] P. Grassberger and I. Procaccia. Measuring the strangeness of strange attractors. *Physica*, 9D:189–208, 1984.

[11] P. Grassberger, T. Schreiber, and C. Schaffrath. Nonlinear time sequence analysis. *Int. J. Bifurcation and Chaos*, 1:521–547, 1991.

[12] L. M. Hively, N. E. Clapp, C. S. Daw, W. F. Lawkins, and M. L. Eisenstadt. Nonlinear analysis of eeg for epileptic seizures. Technical Report ORNL/TM-12961, Oak Ridge National Laboratory, Oak Ridge, TN 37831, 1995.

[13] W. F. Lawkins, C. S. Daw, D. J. Downing, and N. E. Clapp. The role of low-pass filtering in the process of attractor reconstruction from experimental chaotic time series. *Physical Review E*, 47:2520–2535, 1993.

[14] J. Neter, W. Wasserman, and M. H. Kutner. *Applied Linear Regression Models.* Richard D. Irwin, Inc., Homewood, Illinois, 1983.

[15] W.S. Pritchard and D.W. Duke. Measuring chaos in thr brain: A tutorial review of nonlinear dynamic eeg analysis. *Int. J. Neurosci.*, 67:31–80, 1992.

[16] Lawrence R. Rabiner and Bernard Gold. *Theory and Application of Digital Signal Processing.* Prentice-Hall, Inc, Englewood Cliffs, New Jersey, 1975.

[17] S.A.R.B. Rombouts, R.W.M. Keunen, and C.J. Stam. Investigation of nonlinear structure in multichannel eeg. *Phys. Lett. A*, 202:352–358, 1995.

[18] D. Ruelle. *Chaotic Evolution and Strange Attractors.* Cambridge University Press, Cambridge, 1989.

[19] C.E. Shannon and W. Weaver. *The Mathematical Theory of Communication.* The University of Illinois Press, Urbana,Illinois, 1949.

[20] F. Takens. Detecting strange attractors in turbulence. In D. A. Rand and L.S. Young, editors, *Proceedings of the Warwick Symposium*, page 366, New York, 1981. Springer.

[21] J. Theiler. On the evidence of low-dimensional chaos in an epileptic electroencephalogram. *Phys. Lett. A*, 196:335–341, 1995.

INTERNAL DISTRIBUTION

1. M. A. Akerman
2. C. K. Bayne
3. J. Barhen
4. M. D. Cheng
5–9. N. E. Clapp,Jr.
10. T. S. Darland
11–15. C. S. Daw
16. D. J. Downing
17. L. J. Gray
18. W. C. Grimmell
19–23. L. M. Hively
24. P. Kanciruk
25. M. R. Kennel
26–30. W. F. Lawkins
31. M. R. Leuze
32. R. C. Mann
33. M. D. Morris
34. C. E. Oliver
35. G. Ostrouchov
36–40. V. Protopopescu
41–45. S. A. Raby
46. R. F. Sincovec
47. E. C. Uberbacher
48. K-25 Applied Tech. Library
49. Y-12 Technical Library
50. Laboratory Records - RC
51–52. Laboratory Records Dept.
53. Central Research Library
54. ORNL Patent Office

EXTERNAL DISTRIBUTION

55. Dr. Henry Abarbanel, Institute for Nonlinear Science, University of California, San Diego, La Jolla, CA 92093-0402

56. Prof. Cor M. van den Bleek, Dept. of Chemical Process Technology, Delft University of Technology, P.O. Box 5045, 2600 GA Delft, THE NETHERLANDS

57–61. Dr. Michael Eisenstadt, Knoxville Neurology Clinic, 930 Emerald Avenue, Suite 815, Knoxville, TN 37917

62. Dr. Celso Grebogi, Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742

63. Dr. Dan Hitchcock, Office of Scientific Computing, ER-7, Applied Mathematical Sciences, Office of Energy Research, U. S. Department of Energy, Washington, DC 20585

64. Dr. Fred Howes, Office of Scientific Computing, ER-7, Applied Mathematical Sciences, Office of Energy Research, U. S. Department of Energy, Washington, DC 20585

65. Dr. Gary Johnson, Office of Scientific Computing, ER-7, Applied Mathematical Sciences, Office of Energy Research, U. S. Department of Energy, Washington, DC 20585

66. Dr. A. M. Liebetrau, Computational Sciences Department, Battelle-Northwest, P. O. Box 999, Richland, WA 99352

67. Dr. J. Lijengren, Pacific Northwest Laboratories, P. O. Box 999, Richland, WA 99352

68. Dr. Oscar Manley, Division of Engineering, Mathematical and Geosciences, ER-15, Office of Basic Energy Sciences, U. S. Department of Energy, Washington, DC 20585

69. Dr. David Nelson, Director of Scientific Computing, ER-7, Applied Mathematical Sciences, Office of Energy Research, U. S. Department of Energy, Washington, DC 20585

70. Dr. Douglas W. Nychka, Statistics Department, North Carolina State University, P.O. Box 8203, Raleigh, NC 27695-8203

71. Mr. Brent Pulsipher, Computational Sciences Department, Battelle Northwest, P. O. Box 999, K1-86, Richland, WA 99352

72. Dr. Jaap C. Schouten, Dept. of Chemical Process Technology, Delft University of Technology, P.O. Box 5045, 2600 GA Delft, THE NETHERLANDS

73. Dr. Alan Solomon, P. O. Box 227, Omer 84965, Israel

74. Dr. William Thomas, University of Tennessee Veterinary School, Department of Small Animal Clinical Sciences, P. O. Box 1071, Knoxville, TN 37901-1071

75. Office of Assistant Manager for Energy Research and Development, Deaprtment of Energy, Oak Ridge Operations Office, P. O. Box 2001, Oak Ridge, TN 37831-8600

76–77. Office of Scientific and Technical Information, P. O. Box 62, Oak Ridge, TN 37830