Detecting non-linearities in neuro-electrical signals:
a study of synchronous local field potentials

Johannes Müller-Gerking [a], Jacques Martinerie [a], Sergio Neuenschwander [b],
Laurent Pezard [a], Bernard Renault [a], Francisco J. Varela [a]

[a] Unité de Psychophysiologie Cognitive (LENA), CNRS URA 654, Faculté de Medicine Pitié-Salpêtrière, 47 Blvd. de l'Hôpital,
75651 Paris Cedex 13, France
[b] Max Planck Institut für Hirnforschung, Deutschordenstraße 46, 60528 Frankfurt am Main, Germany

Abstract

The question of the presence and detection of non-linear dynamics and possibly low-dimensional chaos in the brain is still an open question, with recent results indicating that initial claims for low dimensionality were faulted by incomplete statistical testing. To make some progress on this question, our approach was to use stringent data analysis of precisely controlled and behaviorally significant neuroelectric data. There are strong indications that functional brain activity is correlated with synchronous local field potentials. We examine here such synchronous episodes in data recorded from the visual system of behaving cats and pigeons. Our purpose was to examine under these ideal conditions whether the time series showed any evidence of non-linearity concommitantly with the arising of synchrony. To test for non-linearity we have used surrogate sets for non-linear forecasting, the false nearest strands method, and an examination of deterministic vs stochastic modeling. Our results indicate that the time series under examination do show evidence for traces of non-linear dynamics but weakly, since they are not robust under changes of parameters. We conclude that low-dimensional chaos is unlikely to be found in the brain, and that a robust detection and characterization of higher-dimensional non-linear dynamics is beyond the reach of current analytical tools.

1. Introduction

1.1. Non-linear dynamics and the brain: The question

The main purpose of this paper is to report on a first systematic attempt to statistically prove the presence of non-linearities in time series of single-cell neuroelectrical data recorded under carefully controlled and behaviorally significant conditions. This is a much less ambitious task than proving the presence of low-dimensional chaotic attractors in lesser controlled and more macroscopic signals in the brain (such as scalp EEG recordings). Checking for non-linearities is a modest but necessary step towards a deeper understanding of the dynamics. Should the detection of non-linearities in this test case prove to be impossible, the search

for chaos in other conditions of brain studies would be seriously put into question. As detailed in this paper, non-linearities in our signals appear to be present, but rather weakly.

In fact, the study of brain signals with the methods of non-linear dynamical systems theory has gained much attention during the last decade [3]. Most effort has been put in estimating dimensions of putative low-dimensional chaotic attractors in scalp EEG recordings following the Grassberger–Procaccia method [10] or variants thereof. Such early approaches are known today to be poor means to identify low-dimensionality, and notoriously unreliable since these algorithms may erroneously claim low-dimensionality where the signal is actually linearly correlated (colored) noise [30], especially when applied to short time series of non-stationary systems. Most of the results obtained with dimension methods have to be reviewed with great care. Recent results bear out this conclusion in EEG [45,41,32].

In addition to problems stemming from this type of signal analysis, the use of EEG scalp recordings, which are large spatial averages of neuronal masses filtered by the skull and the scalp, recorded either free-running or during complex cognitive tasks (such as counting or semantic discrimination), preclude very clear interpretations of a quantitative analysis. It is hard to imagine that the brain enters into a very homogeneous, stationary low-dimensional "counting-numbers" dynamics. Tasks of this kind involve multiple-distributed and temporally unstable mechanisms: changes in dimensions in terms of overall cognitive processes of the brain is bound to be an impossible task.

In order to get more telling results, we have followed here a more simple and precise strategy. Firstly, we have selected our time series from very local neural responses. More precisely, we have studied recordings from two micro-electrodes placed in brain areas involved in the early stages of visual pathways (the optical tectum of pigeons and the thalamus of cats). The electrodes were such that they recorded from a few neighboring neurons, as seen in their local summed electrical field or local field potential (LFP). Further, the animals were recorded while fully awake and subject to visual stimulations which are perfectly well controlled.

Second, we are interested in the data that display transiently oscillatory behavior, leading to a coherent or synchronous activity among multiple neurons. It has been argued that these neural synchronies are a good correlate of cognitive processes since they permit an ensemble of neurons to be coactivated to produce a behavioral meaningful action (see e.g. [36]) Our study is focused on those periods of neural activity where synchronies have been observed during a behavioral visual discrimination [22,20]. We can be maximally assured that our time series do reflect locally a dynamics of biological relevance.

Phase synchronization of brain signals can be understood as a complex, emerging phenomenon between distant neuronal ensembles. We expect therefore to find traces of non-linear dynamics in association with the synchronization process which our data sets follows very precisely. Of course, transition states in networks of coupled non-linear oscillators are well known in some model cases [38,15], and this constitutes a further justification for the search of traces of non-linearities by statistical methods based on dynamical systems theory. Only through a multi-approach testing can we be assured that such non-linearities are present in the neuroelectric data. More specifically, we have applied three methods to test for non-linearity: the construction of phase-shuffled surrogate data and subsequent comparison of their forecastability by a simple non-linear model, the method of "false nearest strands" and the "deterministic-versus-stochastic modeling" procedure.

This paper is structured as follows: for the remainder of this section we clarify our overall setting and notation; in Section 2 we detail the three methods employed in this study. Section 3 presents our findings, followed by Section 4 where we draw our tentative conclusions and suggestions for further improvement in the study of non-linear properties of brain data.

1.2. Dynamics and determinism: Formal setting

As is well known, non-linear models may be very simple and still account for a large part of very complex behaviors. Unfortunately, however, in the first years of enthusiasm researchers often forgot that while some low-dimensional models may behave in a complex unpredictable fashion, there is nothing indicating that all random looking behavior is due to low-dimensional chaotic systems. If one claims on philosophical grounds that all events are deterministic, as long as the systems ruling these deterministic events are of extremely high dimension, the claim does not help in any way the analysis of data from such systems. Such systems will appear indistinguishable from random noise, and this is why for decades statistical methods have worked out well and are still valid approaches to these kinds of observations.

The novelty opened by non-linear dynamical methods lay in the middle ground between these extremely high-dimensional systems and the simplicity of low-dimensional (deterministic) linear systems. So, the only systems where the newer findings may find application are still systems of relatively *low* dimension. While chaos (i.e., random looking behavior from deterministic systems) seems to be a universal phenomenon in nature, low-dimensional chaos may be much less frequent. How "low" the dimension must be in order to see the determinism depends on the system itself (e.g., its Lyapunov exponents) and the observation conditions. From scalar time series of moderate length, the dimension should be sensibly less than ten, to distinguish the data from purely random ones. In the following, when we write "chaos", we implicitly mean low-dimensional dynamics. Then there are "mixed" systems, which, under lucky circumstances, may be separated into a dominant low-dimensional part and high-dimensional contributions. The issue of concern to us can now be stated thus: within the panorama just sketched, to which category do locally recorded neuroelectrical responses from a behaving animal belong?

To introduce our notation, we follow usage by assuming when analyzing data recorded from an unknown dynamical system, such as the brain, that the system is defined in terms of (unknown) ordinary differential equations,

$$\frac{du(t)}{dt} = F(u(t)) + \eta(t),\tag{1}$$

with $u \in \mathbb{R}^n$ and $F : \mathbb{R}^n \to \mathbb{R}^n$ a "well behaved" function. In the absence of $\eta(t)$, that is, $\eta(t) \equiv 0$, Eq. (1) implicitly defines a deterministic flow in the state space $\mathbb{R}^n$; a solution of Eq. (1) $u(t)$ for some time interval $t_0 \ldots t_1$ describes a trajectory in the state space (or phase space) $\mathbb{R}^n$.

The forces $\eta(t)$ reflect our incomplete knowledge of the state of the whole system. In simulation studies $\eta(t)$ is usually considered as noise and modeled by an appropriate stochastic process. In the parlance of Tong [46] $F(u(t))$ is called the (deterministic) *skeleton* of the dynamics which is *clothed* by the noise. Since this term enters the dynamics of the system, it is referred to as *dynamical noise*. We do not want to enter the eternal debate whether there exists true randomness or not: $\eta(t)$ may result from perfectly deterministic processes, from which, however, we have no information in our data. Given the multitude of possible external influences, we expect these processes to be of extremely high dimension, so that they effectively appear as random.

We take now a series of k simultaneous measurements (in our case electrical recordings) of our system (1) at discrete time steps Δt and obtain a time series of N points,

$$\{x_{t \equiv i \cdot \Delta t}\} \quad x \in \mathbb{R}^k \quad 0 \leq i < N.\tag{2}$$

in observation space $\mathbb{R}^k$, with $1/\Delta t$ the *sampling frequency*. This observation process may be modeled mathematically by

$$x_t = h(u(t)) + \theta_t,\tag{3}$$

with some (unknown) observation function $h : \mathbb{R}^n \to \mathbb{R}^k$ and observational noise θ_t. Again, we want to stress the point that what we call "observational noise" may well stem from other deterministic processes, from which we have no further knowledge, and which are of very high dimension. In this paper we consider only scalar time series, that is, $k = 1$ and we write the series as $\{x_t\}$. As we shall see, multi-site recordings will most certainly be essential for future studies of brain dynamics.

1.3. Reconstruction of the system's state space

Since we don't have access to the dynamical system itself but only to a time series of observations, we have to reconstruct a suitable state space for the unknown system from these observations in order to find any characteristics of the underlying dynamics. This is possible by the method of *time delays*. Sauer et al. [35] extending on previous work of Whitney [49], Packard et al. [25] and Takens [39] proved that in the absence of noise, unlimited number of data points and fairly generic conditions on the system equations F and observation function h, the *delay vectors* $x(t) \in \mathbb{R}^d$, constructed according to the rule

$$x(t) = \left\{ x(t), x(t - \tau), x(t - 2\tau), \ldots x(t - (d - 1)\tau) \right\}, \tag{4}$$

form an embedding of the dynamical system. The condition is that the *embedding dimension d* be greater than two times the box-counting dimension of the attractor A of the dynamical system. Then, the map from the attractor A to time-lag space $\mathbb{R}^d$ defined by this prescription is one-to-one on A and an immersion on each compact subset C of a smooth manifold contained in A. If the reconstruction is an embedding, the most important properties of the original system are preserved in the reconstructed space and characterizations of the reconstructed system apply equally well to the original one. It should be noted that the use of time delay vectors is by no means new and goes back at least to the work of Yule [50] (see [6]). The truly new insights brought by dynamical systems theory is the existence of geometrical invariants which are preserved under the reconstruction.

Other approaches to reconstruct a system's state space are possible and have been proposed (see e.g. [4,9]). These methods start from the time-delay reconstruction and try to compress the information given into a new space of smaller dimension, and subsequent calculations are less time consuming. In this paper, we have not tried to optimize over the direct time-delay approach.

There is a vast literature about how to choose the "best" embedding parameters delay time τ and embedding dimension d for the time lag method (see e.g. [9,18,33] and references therein). While the general theorems do not help in the choice of these parameters, the concrete values are of great importance for practical state space reconstruction, that is, reconstruction from a limited set of noisy observations. Again, we did not seek to optimize our choice but tried systematically over a wide range of "reasonable" parameter values according to the intrinsic time scales of our data. We do this also in order to see the robustness of the non-linearity: non-linearities stemming from low-dimensional dynamics should be insensitive to reconstruction parameters in a wide range. The discrimination statistics should improve for higher embedding dimensions. Finally, the method of false nearest strands we use to test for non-linearity, has originally been designed for the purpose of estimating these embedding parameters.

The presence of noise in the dynamical system or in the observations seriously complicates the situation. Casdagli et al. [6] worked on the problem of information lost due to the projection to a low-dimensional observation space by the measurement function. In this view, each measurement is considered as carrying information about the localization of the state $u(t)$ in the original state space. Uncertainties or noise in the observations allow only for approximative localization of the original state. Casdagli et al. show that this localization may be impossible from scalar measurements, even for moderate values of dimension, Lyapunov

exponent and observation noise. Consequently, any trace of determinism is lost and the observation of the chaotic dynamics becomes a truly random process. The only way to "see" more of the determinism is to gather more and better information from the system. This includes as well knowledge on the system other than from its observation, possibly from some explicit model.

1.4. The search for non-linearities

Characterizing a dynamical system on the basis of an observed time series consists of two fundamental steps: the reconstruction of a suitable state space, and the estimation of some characteristic quantities of the dynamics, estimated on the trajectory in reconstructed space. Since we do not have any idea of the system's dynamical equations and the analytical form of the observation function we have no direct way to check whether our embedding procedure succeeded and that we actually have a relevant representation of the dynamics. What usually has been done in this case is to calculate some system invariant (such as correlation dimension or Lyapunov exponents) as a function of increasing embedding dimension. Once a sufficient embedding dimension reached, the estimated invariant should saturate at the correct value. High-dimensional systems should never show saturation, independent of embedding dimension.

This approach, however, has proven to be of limited reliability. Procedures like the Grassberger-Procaccia algorithm [10] involve subjective judgments, for example, on the presence of plateaus which make its application difficult. Even if errors like those reported by Osborne and Provenzale [23] can be avoided when carefully applying the method [40,11], calculations of some invariant alone are no more accepted as reliable indicators of low-dimensional dynamics.

Rather, the idea that has been explored recently is to *test* the data first for the presence of non-linearities: one tries to reject the null hypothesis that the signals recorded are linearly correlated Gaussian noise. From the viewpoint that the time series recorded contains *information* about the dynamical system under investigation, testing for non-linearities can be regarded as a test whether the time series recorded contains information *beyond* the classical second order statistics, that is, auto-correlation or power spectrum. Once the presence of a non-linearity is clearly proven, one may attempt to characterize it.

We have mentioned above how information about the dynamical system can get lost by the projection to a low-dimensional observation space. Inferring from reconstructed trajectories introduce additional sources of estimation errors, due to insufficient length of the time series, or non-stationarities. In line with the Central Limit behavior, this information lost due to low-dimensional projections, uncertainties and estimation errors, will eventually add up in a way such that the time series appears to be Gaussian noise (perhaps transformed by a static non-linearity in the measurement function).

Technically speaking, we try to reject the hypothesis that linear models of the form

$$x_{t+1} = \sum_{i=0}^{p} \alpha_i x_{t-i} + \sum_{i=0}^{q} \beta_i \epsilon_{t-i} \tag{5}$$

(linear auto-regressive, moving average (ARMA) models), with real parameters α_i, β_i and Gaussian innovations ϵ_t is all we are able to infer from the given data. These time series are completely determined by their spectra, or, because of the Wiener-Kinchin theorem, by their auto-covariance. No additional information is specified, for example, on the relationships between the phases of different independent frequencies. Non-linearities, by contrast, necessarily involve specific relations among the phases of different frequencies. This is why filtering out some frequency band to look for chaos in brain recordings (e.g. the alpha band 8-12 Hz [37]) can never succeed since by the filtering we eliminate any information which might allow us to detect the non-linearity.

There are many approaches to test for the presence of non-linearities in time series. Classical statistical methods first "bleach" the data by fitting linear auto-regressive models (AR), eliminating thus all linear correlations. Then, one tries to find some additional structure in the residuals of the linear fit, by using models with higher order terms [16,17,34,27,47]. These methods are very well suited for data whose linear part can be modeled by low-order AR models and test the hypothesis of linearity against rather specific alternative non-linear models, like higher order Volterra expansions [17], smooth threshold auto-regressive models [47,16], or smooth exponential threshold auto-regressive models [27]. The assumptions of such alternative hypotheses demands supplementary knowledge about the putative processes which have created the data, a knowledge which is not available to us in the case of brain time series. Also, these tests are very sensitive to incomplete elimination of the linearities, leading to false claims for non-linearity where some linear relationship is still present in the residuals. Thus, highly auto-correlated data like ours are difficult to treat with these algorithms. An exception is provided by Hinich, who proposed a test for non-linearity based on higher order spectral components [12], in particular the estimation of the bispectrum, therefore searching for inter-frequency relationships which are typical for non-linearities.

Consequently, we propose to concentrate here on another series of methods inspired by non-linear dynamical systems. The objective is the detection of traces of some dynamical non-linear system in the recorded data. The results in this paper have been obtained with three methods: (1) First, we compare phase-randomized surrogate data with the original time series by means of a simple non-linear forecasting procedure. (2) Second we apply the recent method of "false nearest strands" which probes into some geometrical properties of the reconstructed state space. (3) Third, we use a method which tries to evaluate the degree of determinism, as measured by out-of-sample forecasts of a continuum of model classes. We preferred these tests since we hoped to gain further insight into the nature of the non-linearities. In particular, data from low-dimensional chaotic systems should show a characteristic behavior in function of the embedding parameters. The more classical statistical approaches mentioned above do not provide this distinction. The following section explains these methods in greater detail.

2. Methods

2.1. Methods for time series analysis

2.1.1. The method of surrogate data

The method of surrogate data is an application of so-called statistical *bootstrap* methods (for a review see e.g. [8]). These methods allow the testing of statistics against null hypotheses whose distribution function cannot be derived analytically. Instead, the distribution of the statistic of interest under some H_0 is estimated by computer simulations and the test is done against this empirical distribution. The main difficulty is the procedure to create convenient realizations under H_0, which is necessary for reliable estimation of the statistics' distribution. In our case, we want to test against a H_0 of linearly correlated noise. Therefore we need an algorithm which creates many realizations with the same linear properties as our original data but otherwise random. Usually this is done by randomizing the phases in Fourier space. This preserves the frequency spectrum, while destroying any relationship which might exist among the phases and which would account for non-linearities in the tested data [1] . Surrogate creation by this technique was introduced in the field of non-linear data analysis by Osborne

---

[1] Due to the central-limit behavior and the Gaussian distribution of the random numbers used to shuffle the phases, the individual points of the surrogates time series created as described, have a tendency towards a Gaussian distribution themselves. This will give false positive results if the distribution of the original time series had a strongly non-Gaussian distribution. To protect against this case one modifies the

et al. [24] and made popular by Theiler et al. [42,43]. A parallel implementation has been proposed by Kennel and Isabelle [14].

In order to discriminate the possibly non-linear time series from its linear surrogates, we need a statistics which depends crucially on the supposed non-linear properties. Several methods have been proposed [42,43,26,14]. Theiler et al. recommend using several discrimination statistics based on approved methods of non-linear data analysis, namely, correlation dimension, non-linear forecasting error and Lyapunov exponents.

The statistics are calculated for the original time series and tested against the distribution estimated from all surrogate time series. If the original data can be distinguished from the re-sampled ones with good confidence we may reject the hypothesis that the data arose from a process described by the linear model and assume non-linear structures to be present. The significance may be estimated as difference of the original and the mean surrogate value, normed by the standard deviation of the surrogate values. Let Q denote the statistics, $\mu_{H_0} = \langle Q_{surr} \rangle$ the mean over Q applied to all surrogate data, and σ_{H_0} the standard deviation of Q_{surr}. Then define S by

$$S = \frac{|Q_{orig} - \mu_{H_0}|}{\sigma_{H_0}}. \tag{6}$$

The units of this dimension-less quantity are usually called "sigmas". If, in addition, the distribution of the statistic is Gaussian, as numerical experiments indicate, a P-value can be calculated by $P_G = 1/2[1 + \text{erf}(S/\sqrt{2})]$.

Alternatively the significance can be estimated by the Monte Carlo probability P_{MC}. This robust empirical measure of probability is defined as

$$P_{MC} = \frac{\text{number of cases } (Q \leq Q_{orig})}{\text{number of cases}}, \tag{7}$$

where "number of cases" includes both surrogates plus the original. Here we are interested in finding Q_{orig} smaller then the statistics for the surrogates, such as non-linear forecasting error or correlation dimension. We can reject the null hypothesis at a confidence level of $P_{MC} = 1/(N_{surr}+1)$ if $Q_{surr} < Q_{orig}$ for all N_{surr} surrogates (see [31]).

Another evaluation of the significance of non-linearity used here, well suited in the case of non-linear forecasting errors as discrimination statistics, has been proposed by Kennel and Isabelle [14]. They use a simple model where the forecasting is done by searching the nearest neighbor in phase space, excluding those data points which are closer in time than some empirically determined decorrelation time, thus correcting against oversampling. Unlike Theiler et al. they don't compare the means of the prediction errors for the original and the surrogate data, but the distributions of the forecasting errors themselves by means of the Mann–Whitney rank sum test. This parameter-free test calculates from the two sets of forecasting errors a quantity Z which is normally distributed with zero mean and unit variance under the null hypothesis that the two samples came from the same distribution. Since this statistical test supposes the samples to be independent, and the forecasting errors of successive points of a highly autocorrelated time series are surely not independent, only a subset of the

---

null hypothesis and tests against the stronger null hypothesis that *the data stem from a linear random process, transformed by a static non-linearity*. To create surrogates for this stronger null hypothesis, one first re-scales the original data to have a Gaussian amplitude distribution. That is, one makes a series of Gaussian random numbers with the same length as the original data and sorts them so that both series have the same rank structure. From this "gaussianized" time series a surrogate is built as described above. The correct surrogate finally is obtained by reordering the original data as to have the same rank structure as the phase randomized version of the gaussianized data. Evidently, this preserves the distribution of the original data points, since the surrogates are nothing but a shuffled variant of them. The trick is a controlled shuffle which preserves the linear dependencies of the original time series. Alternatively, one may work with the gaussianized time series itself.

errors are taken into account. In what follows, the surrogate tests were done with a Mann–Whitney statistics, following Kennel, on whose implementation ours is based.

As we discuss below, it seems that the FNS method is both more reliable and less costly in CPU time than the surrogate method just described. However, in order to provide a clear link with current literature which uses extensively the surrogate approach, we have used both tests for the set of data from the cat's LGN, but not for the pigeon tectum data. The LGN data was more adapted, since it was less oversampled, and better stationarity allowed for longer segments of analysis.

2.1.2. The method of false nearest strands

The initial motivation of this method is to find best embedding parameters, time-lag τ and dimension d. It can, however, easily be extended in a natural way to provide a fast and reliable test for non-linearity. In order to test if some embedding makes a deterministic map, one examines the number of self-crossings of trajectories in the reconstructed state space. If the unknown evolution in the original state space is deterministic and the embedding is "good", that is, one-to-one on the attractor, there will be no self-crossings of trajectories in the reconstructed state space. The method of *false nearest strands* [13] improves over a former approach of *false nearest neighbors* [2] in that it accounts for highly oversampled data and small time delays.

The idea behind *false nearest neighbors* is to check whether nearest neighbors in the reconstructed state space are neighbors due to the dynamics ("true nearest neighbors") or rather due to the projection of the original state space into a space of inappropriately low dimension ("false nearest neighbors"). To this end one looks for nearest neighbors in a (d)-dimensional reconstruction space and calculates the distance of these points in a $(d + 1)$-dimensional space, constructed by adding the $d + 1$st coordinate. If the distance in the higher dimensional space is very large, we have found what is called a *false neighbor* since the two points are close in d dimensions only due to the too low-dimensional projection. In contrast, *true* neighbors in d-dimensional space will remain neighbors in $(d + 1)$-dimensional space. The ratio of false nearest neighbors over all pairs of neighbors tested should go to zero when we have found a good embedding. This statistic based on counts of false nearest neighbors, however, is affected by high temporal correlation in the data due to oversampling, or embeddings with very short time delays. These consideration has led the authors to count *false nearest strands* rather than *false neighbors* individually. A *strand pair* is made of all pairs of points within some time interval which are mutually nearest neighbors and direct iterates of one another. Now, instead of identifying false embeddings by the number of false nearest neighbors, we examine the distance of strand pairs in different embedding dimensions. There are multiple ways of defining when a strand pair is false. Kennel and Abarbanel [13] chose to designate a strand pair as false, if the average "extra distance" in $(d + 1)$-dimensional space is too large. That is, they compute the mean absolute distance on the $d + 1$st coordinate for a whole strand pair. The strand pair is declared false if this additional distance is larger than some parameter p of order unity, times the natural radius of the attractor, R_A. The final statistic is then the ratio of false nearest strand pairs to the total number of strand pairs.

To summarize briefly (see [13] for details), the false nearest strand procedure is performed in this way: (1) choose some time lag τ and embedding dimension d and embed the scalar time series with this time delay in a $(d + 1)$-dimensional space. This space is rotated and stretched to normalized principal components, so that the coordinates are now linearly decorrelated. (2) The trajectory in the new, rotated space is projected down to d dimensions by discarding the $d + 1$st coordinate. (3) In the d-dimensional space, search the nearest neighbors to all points, excluding those whose time index is closer than some decorrelation time, usually in the order of two times the autocorrelation time. (4) From this list of nearest neighbors establish the list of strand pairs. In function of the $d + 1$st coordinate decide for each strand pair if it is a true or false one.

We have also compared the ratio of false strands with that obtained by randomly shuffling the last component. If the signal examined stems from a truly low-dimensional dynamics, the first ratio should decrease with higher embedding dimension (and finally go to zero for noise free data) while the ratio of false strands with random last component should remain at a high level around 50%, independent of the embedding parameters. With reasonably clean low-dimensional data the difference between original data and shuffled last component should be visible by eye. This is why the authors did not perform any supplementary statistical evaluation of the strand ratios. For our data, the difference cannot be evaluated as easily. Therefore we added in our implementation the obvious Monte Carlo simulation, performing numerous shufflings (at almost zero cost) and estimating the probability of linearly correlated noise by the statistics based on Gaussian distribution equation (6) and the empirical Monte Carlo probability equation (7) as stated above.

In this article we have favored the FNS method over the surrogate construction for two main reasons. First, it does not require the time consuming search for nearest neighbors for each surrogate which is done only once in the FNS method. Secondly, the surrogate construction procedure may introduce spurious artifacts (see for example [44]).

2.1.3. Deterministic versus stochastic modeling

The "deterministic vs stochastic" (DVS) method [7,5] consists of fitting a family of local linear models to the data and analyzing their prediction accuracy in function of various parameters which determine the model class. The procedure evaluates the forecasting performance of models ranging from local linear, that is, globally non-linear to global linear ones. Local linear models perform better on chaotic data than global linear ones, data from linear correlated noisy devices are better modeled by globally linear models. Data from low-dimensional noisy chaotic devices will show lowest forecasting errors with intermediated models.

The system's state space is reconstructed by the method of time delay embedding to obtain trajectory points as discussed above. One chooses test vectors x_i and fits affine models of the form

$$x_{j+T} \approx f^T(x_j) \tag{8}$$

to selected vectors out of a training set of N_F trajectory points preceding the test vector in time. The model is then used to estimate a T step ahead prediction of the test point by

$$\hat{x}_{i+T} = f^T(x_i) . \tag{9}$$

We select the k nearest neighbors to the test vector in state space as basis of the fitting model. Model parameters $\alpha_i, i = 0, \ldots, d$ are determined as to minimize the squared error of models,

$$x_{j(l)+T} = \sum_{n=0}^{d-1} \alpha_n x_{j(l)-nr} + \alpha_d , \quad l = 1, \ldots, k . \tag{10}$$

Since there are $(d + 1)$ parameters, the minimal k to solve this equation is $k = d + 1$. The solution is unique, if the matrix composed of the x_l is non-singular. This is rarely what we want, since we do not expect single data-points determine the future behavior of the trajectory via a linear model. A reasonable lower limit of k to have a fit is two times the mathematical minimum, that is $k = 2(d + 1)$.

By varying the number of nearest neighbors used to build the model in the range of $k = 2(d + 1)$ to $k = N_F$ (that is, the whole fitting set), we sweep the whole model classes from the deterministic extreme (local linear models) to the stochastic extreme (global linear models). Notice that for $k = N_F$ the fitted model is nothing but a linear autoregressive model of order d.

This procedure of model building is repeated for a large number N_T of test vectors, and the mean absolute forecasting error is computed,

$$E_d(k) = \sum_{\text{all test vectors}} \frac{|\hat{x}_{i+T} - x_{i+T}|}{N_T} . \tag{11}$$

Plots of these error curves for different embedding dimensions d against the number of nearest neighbors, k, is called a DVS plot. If models near the deterministic extreme give the most accurate short-term forecasts, then we have strong evidence for low-dimensional chaotic behavior in the data. This can been seen in Fig. 1a where a DVS plot for the chaotic Lorenz system is shown. Once the sufficient embedding dimension (3) has been reached, the forecasting error decreases drastically for local linear models. Global linear (AR) models produce errors about 200 times higher. When we add 25% (RMS) white noise to the same data, we observe first a general degradation of the forecasting (Fig. 1b). The errors go up from 0.002 for the noise-free data to 0.31 with noise added, taken from the respective best models. The signature of the DVS plot is typical for a noisy observation of a low-dimensional chaotic system: we need more observations in order to "average out" the noise, but the local structure in reconstructed state space is still prevalent. Therefore, the best forecasting is obtained with intermediary model classes. Linearly correlated noise (Fig. 1c) shows a further diminution of the forecasting precision, and best prediction at globally linear models. The data for this calculation has been obtained by phase-shuffling the noise-free Lorenz data, used for the two plots before. There is no structure in the reconstructed state space which could be exploited by the simple local linear models used. White noise (Fig. 1d) shows the same signature as linearly correlated noise, but the predictability has been lost. The best forecast is the mean value itself, this is why the AR model of lowest order performs best here.

2.2. Methods for the recording of time series

We have applied the methods described above to local field potentials recorded from micro-electrodes placed in the lateral geniculate nucleus (LGN lamina A1) of mildly anesthetized cats and in the optic tectum of awake pigeons. For full details of the methods see [20,21]; we only summarize here some essential points.

In the visual pathways of all vertebrates the retina connects to the brain via the optic nerve which follows two parallel and major routes: the so-called tecto-fugal and thalamo-fugal pathways. In mammals, the pathway with the largest number of connections is the latter one, connecting optic nerve axons to the thalamus, at the lateral geniculate nucleus (LGN). The second pathway goes to the superior colliculus, and both LGN and colliculus are reciprocally connected with the visual cortex (see e.g. [52]). In other vertebrates (such as fishes and birds) the largest pathway goes to the optic tectum (homologous to the colliculus of mammals, belonging to the tecto-fugal pathway) while the second pathway goes to several disperse points of the thalamic complex (see e.g. [51]).

In our study we have chosen recordings from the two preferred connections from the retina: LGN in cats and tectum in pigeons. These structures are thus quite different anatomically, but in both cases there is evidence of a very active modulation from retinal stimulation in the receptive fields of the neurons recorded with micro-electrodes. Tectal cells respond best to moving stimuli (in our case: light-bars), but only briefly; some thalamic cells, in contrast, may show strong, long lasting responses also to static stimuli.

Further, in both cases an active sensory modulation evokes synchronous activity in pairs of neurons (or small groups of multi-site responses) which, although separated by significant distance (up to a few millimeters), they may nevertheless enter into very precise temporal synchrony in their electrical responses while the visual stimulation is present [20].
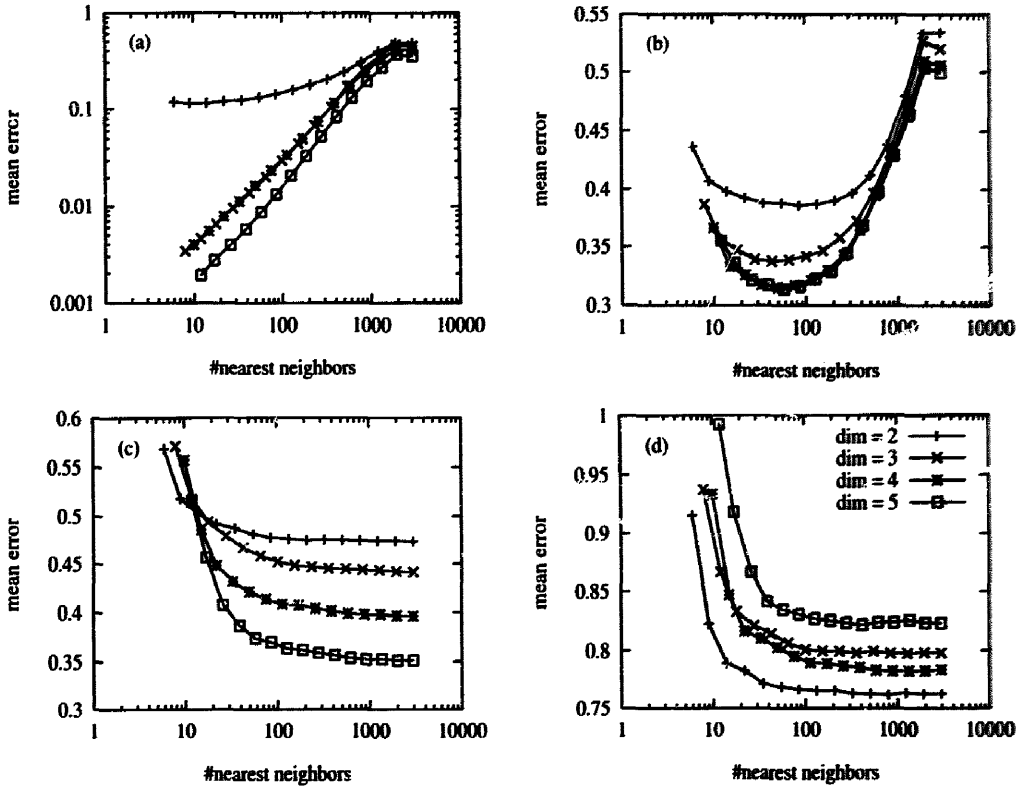
Fig. 1. Examples for DVS plots for embedding dimensions 2 to 5. Horizontal axis: number of nearest neighbors used to build the local linear model. Vertical axis: mean absolute forecasting error. (a) Noise-free Lorenz system. Once a sufficient embedding dimension reached, the forecasting error drops by a factor of almost 200 for m_dels in the deterministic extreme. (b) Lorenz data plus 25% (RMS) white noise. Generally, the forecasting is much worse, the least forecasting errors is obtained for intermediate model classes. (c) Phase randomized Lorenz data as example for linearly correlated noise: least forecasting error for global linear models. (d) White Gaussian noise.

By differential filtering, two types of electrical signals can be recorded from micro-electrodes: spike activity (single or multi-unit) and field potentials. Spike activity can be detected by high-pass filtering of the data in the range of 10 KHz, since they correspond to propagated action potentials of brief (1 msec) duration. Local field potentials (LFP), in contrast, are continuous signals, obtained when filtering up to 100 Hz. Since we record from micro-electrodes, we record potentials only from a local population. It is not precisely known how LFPs are composed from the activities of the neurons around the electrode, but it is widely assumed to be a spatial average over sub-threshold and spike potentials. A *spike-triggered average* of the LFP clearly shows a functional connection between both. Since it is much more convenient to use the continuously sampled LFP instead of the discrete series of spikes, we have worked here exclusively with LFP as our source of data.

The time series studied here correspond, then, to pairs of neurons in either tectum or LGN where an LFP response is detected while a light stimulus is presented twice in front of the receptive field. There are two subsequent periods of response, and each trial is part of a sequence of ten repetitions which are individually analyzed.

3. Results

3.1. Neural responses from the thalamus

3.1.1. General conditions for analysis

We have analyzed 5 sets of local field potential recordings from two electrodes in the lateral geniculate nucleus (LGN lamina A1) of cats (see [21] for details). Each set comprises 10 trials of 10 seconds of continuous recording with a sampling frequency of 1 kHz. There are two experimental conditions: static flashes and moving light bars. For data sets labeled SA and SB stimulation was produced by means of a static light bar, flashed over the receptive field. A trial comprises two stimulations of 2 seconds duration each, preceded and followed by an interstimulus period of 2 seconds. Data sets labeled MA and MB were recorded with stimulation by a single moving light bar across the receptive field, moving forwards and then backwards in each trial. A third data set MM correspond to recordings with stimulations by two moving light bars.

The methods of surrogate data and false nearest strands have been applied to overlapping segments of 2048 ms and 1000 ms length, respectively. To this end, we applied first a low-pass filter with a cutoff (-3 dB point) at 140 Hz to remove high-frequency noise. The filtering has been performed in Fourier domain[2]. Using this a-causal filtering should introduce only negligible side effects, as has been shown in numerical studies [19]. The search for non-linearities has been performed on overlapping segments, cut out from these filtered time series. Each segment is then rescaled to have an approximately Gaussian amplitude distribution before the tests are applied. The time series have low power at frequencies below 5 Hz, so that the signals may be regarded approximately stationary within the segments chosen.

After filtering, the first zero crossing of the auto-correlation function calculated for each of the segments falls around 15-20 ms (equals number of points) for each of the segments. There is very little variation of this characteristic time within the trials. We have chosen therefore the range of embedding time lags as 3 ms to 18 ms in steps of 3 ms. The decorrelation time for the search of nearest neighbors is state space was set constant to 25 ms. The tested embedding dimensions are 1 to 8 in steps of 1. The two methods applied systematically over this range of embedding parameters return the respective statistics for each of these combinations. That is, a $N(0, 1)$ distributed value Z in the case of the surrogate method, and directly an α-value in the case of the FNS method.

We need to consider the problem of deriving inferences from replicated tests. Given some significance level of, say, 95%, we expect 5% of false positive results in these tests. Performing the same test n times evidently increases the absolute number of positive results, true as well as false, and a single statistic above the level of significance is therefore less reliable. To get around that problem, Kennel and Isabelle [14] proposed to raise the significance level accordingly such that the new threshold corresponds to $\alpha_{new} = \alpha/n$.

There is a difficulty with this approach in that we do not perform exactly the same test, but tests with varied parameters. In principle statistics for data from truly low-dimensional dynamics should be insensitive to changes in the embedding in a rather wide range of the parameters, so the assumption of "repeating the same test" is correct. But our data are less conclusive and we proceed therefore as follows: we count the number of times when the statistic reaches the significance level. We require at least 5% out of all repetitions in order to be classified significant. Again, for truly low-dimensional data this should give the same result. But it allows to detect trends in the experimental data even when they are not fully significant. In our case, we perform $8 \times 6 = 48$ tests on each window. A significant non-linearity requires more than or equal ≈ 3 positive results.

---

[2] To do so, we use only 8192 points of the whole time series, starting at 1000 ms.

Table 1

LGN surrogate statistics: counts of segments with more than 0, 2, 4, 5 or 6 individual rejections of H_0 at the 95% level (value $Z < -1.65$) amongst all embedding-sets (48). Total number of segments per data set and electrode: 70. Last column: total count of rejections (out of 3360). Percentage values in parenthesis, rounded.

	#cts>0		#cts>2		#cts>4		#cts>5		#cts>6		total	
NULL	63	(90)	32	(46)	9	(13)	4	(6)	1	(1)	172	(5)
MA-1	66	(94)	46	(66)	24	(34)	17	(24)	12	(17)	272	(8)
MA-2	66	(94)	42	(60)	15	(21)	7	(10)	3	(4)	216	(6)
MB-1	67	(96)	57	(81)	28	(40)	19	(27)	12	(17)	296	(9)
MB-2	63	(90)	34	(49)	11	(16)	4	(6)	1	(1)	188	(6)
SA-1	68	(97)	58	(83)	36	(51)	27	(39)	17	(24)	368	(11)
SA-2	61	(87)	34	(49)	10	(14)	8	(11)	4	(6)	190	(6)
SB-1	66	(94)	35	(50)	15	(21)	10	(14)	8	(11)	220	(7)
SB-2	62	(89)	33	(47)	11	(16)	7	(10)	3	(4)	184	(5)
MM-1	62	(89)	44	(63)	19	(27)	13	(19)	11	(16)	247	(7)
MM-2	63	(90)	40	(57)	13	(19)	8	(11)	4	(6)	211	(6)

In order to verify our strategy, we performed the same tests on a set of "null-trials", constructed by phase-scrambling the data of a "good" set (MA-1) *before* they enter the surrogate or FNS procedures. By this, we have an additional empirical criterion to compare our results.

3.1.2. Results for the surrogate method

Table 1 reports the overall statistics on identified non-linearities in overlapping segments of 2048 ms length, where each trial has been cut into 7 overlapping segments, giving 70 segments for each set and electrode. For each experimental sequence of 10 trials, we count the number of segments (out of 70) where the hypothesis of linearly correlated noise could be rejected at the 95% level, that is, a value of $Z < -1.65$. We report on the cases where H_0 was rejected for more than 0, 2, 4, 5 and 6 amongst the 48 combinations of different embedding dimensions (1–8) and time lags (3–18 ms in steps of 3 ms). We shall refer to one such combination of embedding dimension and time-lag as an *embedding-set*. The scores are listed in columns 1 to 5. In the last column we list the total number of positive embedding-sets (out of $70 \times 48 = 3360$).

We expect 5% of false positives, that is, 168. Indeed, the null-test returns a total of 172 rejections (NULL row, last column), which fits well the theoretical value. Almost all segments for all data sets, true data as well as scrambled data, contain at least one embedding-set where H_0 could be rejected ($\approx 90\%$, first column). At the theoretical significance level (required number > 2, column 2), we detect a rather high ratio of significant non-linearities in the original data (more than 50%). Notice, however, that the counts in the case of the null-test are still very high (46%) and are only slightly different from the true data. The null-test actually falls to the required 5% level for *segments* to be considered significant with more than 5 counts (column 4 shows 6%; 5% out of 70 is 3.5 counts). At this level, there is a clear separation between the true data and the null-test, on the one hand, and between the two electrodes, on the other. Electrode one still counts an enhanced number of significant segments (20–40%), while electrode two shows sensibly less.

In summary, the data clearly displays segments with significant non-linearities. If they stem from a low-dimensional chaotic process, however, they should be insensitive to changes in the embedding parameters in a rather wide range. This is certainly not the case for the present data.

In Fig. 2 (upper row) we plot cumulative counts for individual rejections at $Z < -1.65$ for the ten trials of
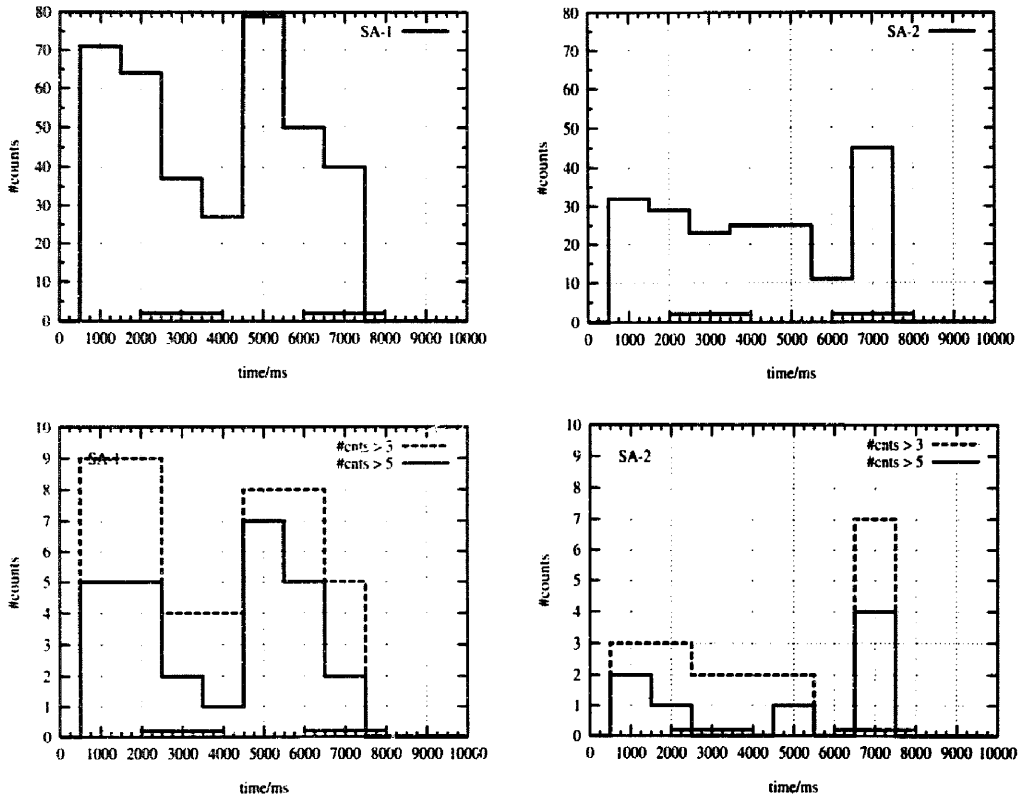
Fig. 2. (top) LGN surrogate cumulative statistics. Results given only for set SA as representative example of all data sets. Plotted are counts of rejections at the individual 95% level ($Z < -1.65$.) Left: electrode 1, right: electrode 2. The high number of counts in set SA-1 around 1600 ms is linked to a highly regular, long lasting oscillation which starts before the stimulation. Stimulation periods are indicated by the bars near the x-axis.
(bottom) Same data as above with the reduced statistics for the surrogate method. Plotted are counts of segments, where the H_0 of correlated noise could be rejected more than 3 times (dotted line) and more than 5 times (continuous line).

data-set SA against time. This is the set which shows the clearest modulation of the occurrence of non-linearities as a function of time. The graphs for the other sets (not shown) have a tendency to be almost constant at a level of 30-40 rejections per time slot.

Fig. 2 (bottom row) displays the results for the same data set, but this time we show the number of segments (out of ten) per time slot, where H_0 could be rejected at the individual 95% level ($Z < -1.65$) for more than 3 and 5 embedding-sets, respectively. The number of significant non-linearities drops drastically when requesting more than the 3 counts that are theoretically necessary. In contrast, the tendency for a coincidence between emergence of non-linearities and stimulation is enhanced in this presentation.

3.1.3. Results for false nearest strands

Table 2 reports the overall statistics on occurrences of non-linearities as detected with the FNS method applied to overlapping segments of 1000 ms length, with 18 segments cut from each trial making a total of 1800 segments analyzed. As in Table 1, the first row reports the results on the null-data set. This test reveals an empirical value of false positives of only 1.3% of the total number of embedding-sets tested (P_{MC} last column;

Table 2
LGN FNS statistics: counts of segments with more than 0–4 individual rejections of H_0 at the 95% level. Values are based on the Monte Carlo probability P_{MC} (percentage values in parenthesis, rounded). Last two columns: total number of rejections for Monte Carlo P_{MC} and Gaussian P_G probability. The results are of the same order.

	#cts>0		#cts>1		#cts>2		#cts>3		#cts>4		P_{MC}	P_G
NULL	78	(43)	25	(14)	8	(4)	2	(1)	1	(1)	114	101
MA-1	87	(48)	33	(18)	10	(6)	4	(2)	1	(1)	135	116
MA-2	87	(48)	32	(18)	13	(7)	5	(3)	1	(1)	138	106
MB-1	77	(43)	25	(14)	7	(4)	3	(2)	1	(1)	116	94
MB-2	94	(52)	44	(24)	18	(10)	4	(2)	1	(1)	163	139
SA-1	79	(44)	32	(18)	15	(8)	6	(3)	6	(3)	159	132
SA-2	84	(47)	31	(17)	9	(5)	4	(2)	2	(1)	134	108
SB-1	89	(49)	39	(22)	19	(11)	4	(2)	1	(1)	152	132
SB-2	95	(53)	38	(21)	15	(8)	2	(1)	0	(0)	150	128
MM-1	87	(48)	34	(19)	15	(8)	6	(3)	1	(1)	143	125
MM-2	96	(53)	32	(18)	7	(4)	0	(0)	0	(0)	135	112

$114/(180 \times 48) = 1.3\%$). This test is therefore more severe that the surrogate test in the previous section. In contrast, the theoretical rejection level for segments (more than two rejections) corresponds to the observed value: 5% out of 180 is 9; the null-test finds 8 (column 3). We also find here that real data and null-set show a similar number of segments with at least one rejection (around 44%, column 1). This percent is almost half as that for the surrogate test.

While about 7% of all segments have been recognized as non-linear at the theoretical significance level (more than 2 counts, column 3), this number decreases drastically when demanding more positive counts than theoretically necessary. We find a persistent non-linearity for the trial series SA-1, which does not change much with higher number of positive counts. This persistency is linked to a very regular, long-lasting oscillation in one of the trials. The superior number of non-linearities in electrode 1 as compared to electrode 2 is less pronounced here or even reversed (sets MA and MB).

The counts presented here are based on the Monte Carlo statistics (Eq. (7)). Counts based on the assumption of a Gaussian distribution are very similar to these. We only report here the total number of rejections (last column in Table 2, P_G) as compared to those found by applying the Monte Carlo statistics (last column, P_{MC}).

The FNS method provides us also with an absolute statistic, the ratio of false nearest strands itself. Remember, a "good" embedding would show no self-crossings of trajectories in the reconstructed phase-space and therefore the ratio of false nearest strands should go to zero for data from deterministic dynamics with sufficiently low dimension and sufficiently high-dimensional embedding space. We selected the segments where H_0 could be rejected individually for 3 different embedding-sets at least, and looked at the absolute FNS ratio. The results are in Table 3. Column 1 lists the minimal FNS ratio found for the selected segments, the second column reports the corresponding FNS ratio with shuffled last component (mean over 39 realizations, see Section 2.1.2). The other columns indicate how often a FNS ratio of less than a certain percent has been found, given that the expected value for correlated noise is around 50%. Although some FNS ratios under 50% are found, they are always far above (40%) the approximately 5% of false strands which the authors of the method judge significant for deterministic dynamics, based on their experience. The last column gives the total number of selected embedding-sets.

In Fig. 3 (upper row) we plot, as before, the cumulative results for the FNS statistics (set SA). In general,

Table 3
LGN: FNS ratios. First column: the least FNS ratio (in %) found in the segments selected as described. The embedding of a deterministic trajectory should give not more than 5% of false nearest strands. Second column: the corresponding mean FNS ratio with shuffled last component. The other columns report how often a FNS ratio of less than 40%,..., 44% has been found. Last column: total number of embedding-sets selected.

	least	shuffled	<40%	<42%	<44%	total
MA-1	42%	50%	0	0	3	15
MA-2	43%	49%	0	0	2	15
MB-1	42%	49%	0	2	4	16
MB-2	39%	50%	1	2	7	25
SA-1	40%	45%	1	5	19	36
SA-2	41%	46%	0	2	8	20
SB-1	43%	47%	0	0	3	26
SB-2	43%	47%	0	0	4	19
MM-1	42%	48%	0	2	5	16
MM-2	42%	48%	0	0	7	12

there is a number of data sets where one does not see any correspondence between non-linearity and stimulation Best cases include the backwards direction of moving bars, and the double moving stimulus (MM). The bottom row in Fig. 3 displays the number of segments (out of ten) for each time slot, where H_0 could be rejected more than 2 and 3 times, respectively. Very few segments are accepted as non-linear, and the relationship to stimulation is not systematic.

3.1.4. DVS results

The DVS method has been applied to all those segments which had more than two positive identifications of non-linearity with the FNS method. First, we search for a total of 104 segments the embedding-sets, where the null hypothesis of linearly correlated noise has been rejected at the 5% level (not more than one embedding-set for each dimension, yielding 200 embedding-sets that have been tested). Then we select each of these particular embedding-sets to build a series of local linear models, varying the number of nearest neighbors from two times the embedding dimension to all fitting vectors in 20 steps (such that the number of nearest neighbors increases exponentially). The data segments having 1000 points, we perform the test on the last 400 points, taking the preceding 600 points (excluding those closer in time than the first zero crossing of the autocorrelation function) as fitting set. The forecasting performance has been evaluated for predictions of 3, 6, 9, and 12 data-points ahead. All segments tested had the signature of high-dimensional systems or linearly correlated noise. In no instance did we find evidence of low-dimensionality by this method.

3.2. Neural responses form the optic tectum

3.2.1. General conditions for analysis

We choose two sets of recordings on two micro-electrodes in the optic tectum of awake pigeons. Each set comprises ten trials of four seconds of recordings at a sampling frequency of 5000 Hz. Stimulation is by a moving light bar over the receptive fields, forwards and backwards in each trial. The neural responses are much more transient than those of the mildly anesthetized cat. This is likely due to the characteristics of tectal neurons, and to the fact of experimenting with a fully behaving animal.
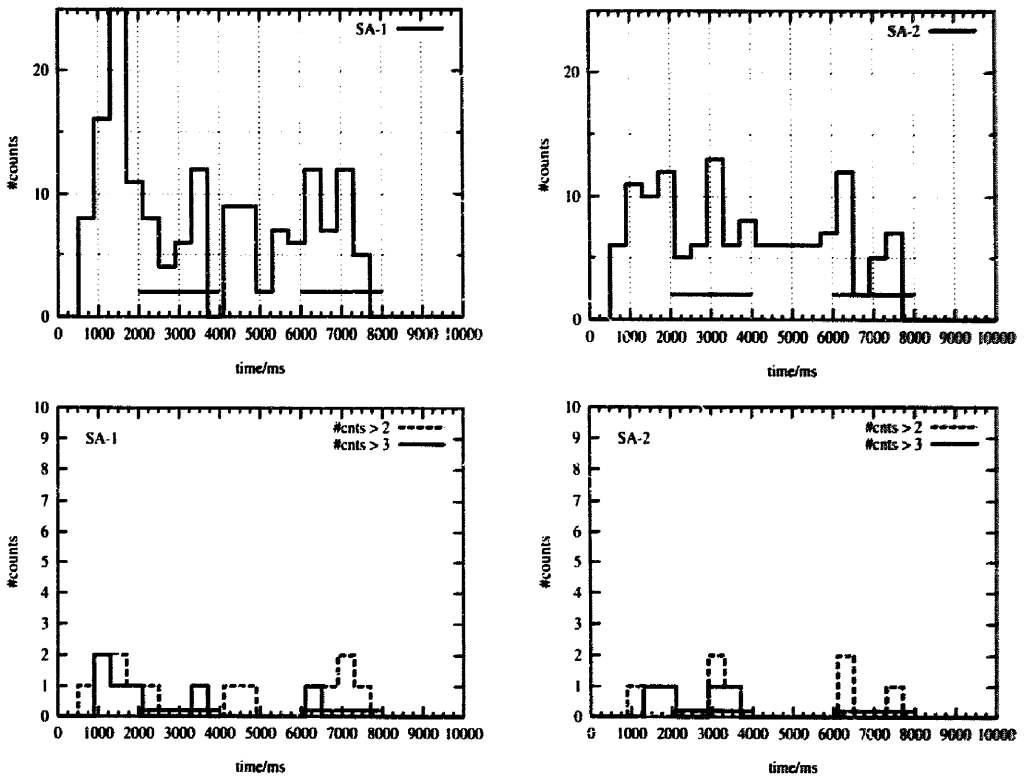
Fig. 3. (top) LGN FNS cumulative results. Same conventions as in Fig. 2; segments of 1000 ms length.
(bottom) Same data as in Fig. 2, reduced statistic for FNS method. Plotted are the number of segments for each time slot, where the H_0 could be rejected more than 2 times (dotted line) and more than 3 times (continuous line).

Set A of the data has been recorded from very closely spaced electrodes at the surface of the tectum (electrode distance: 0.3 mm, depth: 0 μm). The electrodes for set B had a larger separation and record from deeper layers (electrode distance: 1 mm, depth: 100 μm). The light bar in forward direction passes at recording time 1000 ms, and backwards at 2500 ms. Stimulus conditions are the same for both sets of trials. The neural response as revealed by the spike activity is almost immediate, starting at time 1050 ms for all electrodes, and lasting around 200–300 ms. While in the case of the cat the experimental setup and the nature of the data allowed for data segments of 1 to 2 seconds (for the question of reasonable cognitive time scales, see below) segments that long are not possible in the case of the tectum, and we thus choose segments of 400 ms length.

The data are highly oversampled; the original recordings have a high percentage of strictly equal pairs of successive samples. We can therefore reduce the number of data-points in order to reduce computational effort without affecting the information contents. We have down-sampled the data by a factor of two, obtaining one new point from the mean of two successive points. The effective sampling frequency is now 2500 Hz. still very high. This leads to a high risk of artifacts for the surrogate method, which is not the case for the FNS method. As discussed above (Section 2.1.2) we thus decided not to analyze the pigeon tectum data with the surrogate method. In contrast, we have studied the first-differenced data[3] since this considerably enhances stationarity.

---

[3] Each point is calculated as the difference of two succeeding points: $x'_t = x_{t+1} - x_t$.

Table 4
Tectum FNS statistics: counts of segments with more than 0, 3, 7, or 10 individual rejections of H_0 at the 95% level (percentage values in parenthesis, rounded). The last two columns list the total number of rejections (out of 14720), the left column based on Monte Carlo probability, the right column based on the assumption of a Gaussian distribution of the surrogate statistics.

	#cts>0		#cts>3		#cts>7		#cts>10		P_{MC}	P_G
Null	182	(79)	59	(26)	14	(6)	4	(2)	577	506
A1	179	(78)	77	(33)	20	(9)	9	(4)	707	636
A2	187	(81)	77	(33)	25	(11)	8	(3)	776	715
B1	188	(82)	82	(36)	28	(12)	11	(5)	770	677
B2	205	(89)	101	(44)	35	(15)	14	(6)	903	807

This was not necessary for the more stationary cat LGN data.

Evidence for the non-stationarity was clear from the high variability of the first zero crossing of the auto-correlation function as compared to the case of the LGN. For most segments the first zero crossing lies between 20 and 30 points (8–12 ms), some however going up to 60–90 points (24–36 ms). We choose therefore embedding time-lags of 4 to 32 points in steps of 4 points for embedding dimensions 1 to 8. The zero-crossings for first-differenced data (after down-sampling by 2) lie between 9 and 12 points; very few segments had their first crossing at up to 16 points. Time lags in this case has been chosen as 2 to 12 points, in steps of 2 points.

3.2.2. Results for false nearest strands

The FNS method has been applied to data segments of 400 ms (1000 points) length; each trial has been cut in 23 overlapping segments. As before, Table 4 lists the number of segments for the four sets of trials, where individually the null hypothesis of correlated noise has been rejected more often than 0, 3, 7 or 10 times out of all 64 embedding-sets. The percentage values (in parenthesis) are calculated on a total number of 230 segments per set. The last two columns lists the total number of rejections (out of $230 \times 64 = 14720$) for counts based on Monte Carlo Statistics (P_{MC}) and on a Gaussian approximation (P_G). The first row again reports the same results for a series of null-data, created from set B1 by phase-randomizing each segment before it enters the test. The observed count of false positives on the total is 3.9% (P_{MC}, 577/14720).

With a significance level of 95% for each test individually, and 64 different embedding-sets per window tested, we require at least 3 positives (5% of 64) to reject the null hypothesis for that data segment. However, this required 5% level in the null-test is only reached for more than 7 counts (column 3). The number of segments with more than the formally necessary number of individual rejections decreases less markedly than in the case of the thalamus, leaving us with around 10–15% of all segments at the observed 95% level.

If we consider now the ratio of false nearest strands, we find a more convincing result than in the case of the cat. As before, we selected the segments where H_0 could be rejected individually for 5 different embedding-sets at least, and looked at the absolute FNS ratio. The results listed in Table 5 show that we are still far from the 5% for deterministic dynamics, but the best false nearest strand ratios decrease now down to 35% and there is a sizable number of segments between 36% and 40%.

We also applied the FNS calculation to the first differenced data as shown in Tables 6 and 7, which should be compared to Tables 4 and 5. Since we performed the test on only 48 different embedding-sets, the significance level is reached at more than 2 rejections, theoretically, but more than 3 rejections are actually required (column 3). The absolute FNS ratios of the first differenced data are worse than in the case of the raw data. This is to be expected, since first differencing augments noise. However, it is useful to do this calculation to check against the hypothesis that non-linearities arise only because of non-stationarity. Also, for

Table 5

Tectum: FNS ratios. As in Table 3. First column: the least FNS ratio (in %) found in the segments selected as described. Second column: mean ratio of false strands with shuffled last component for this window. The other columns report how often a FNS ratio of less than 36%,..., 44% has been found. Last column: total of embedding-sets considered.

	least	shuffled	<36%	<38%	<40%	<42%	<44%	total
A1	35%	48%	1	5	19	43	109	223
A2	36%	46%	0	6	19	51	133	237
B1	35%	48%	1	6	20	65	133	254
B2	36%	48%	0	8	29	79	181	321

Table 6

Tectum FNS statistics: first differenced data. Compare to Table 4.

	#cts>0		#cts>2		#cts>3		#cts>5		P_{MC}	P_G
Null	134	(58)	26	(11)	12	(5)	2	(1)	248	200
A1	172	(75)	57	(25)	33	(14)	10	(4)	399	334
A2	179	(78)	50	(22)	26	(11)	6	(3)	383	312
B1	161	(70)	60	(26)	42	(18)	15	(7)	447	401
B2	159	(69)	49	(21)	20	(9)	2	(1)	333	278

Table 7

Tectum: FNS ratios, first differenced data. Compare to Table 5.

	least	mean	<38%	<40%	<42%	<44%	total
A1	40%	48%	0	2	9	28	46
A2	42%	49%	0	0	0	11	42
B1	39%	46%	0	1	6	34	91
B2	40%	48%	0	0	2	8	18

truly low-dimensional dynamics first differencing should not change the results [29].

Fig. 4 (top row) shows the cumulative results based on the Monte Carlo estimator of the FNS statistics against time as in Fig. 3 for one of the four sets analyzed. The left column represents the results obtained from the raw data, the right column those of the first differenced data. This set (B1) shows enhanced occurrence of non-linearities during the two stimulation periods (indicated by the horizontal bar) for both raw data and first differenced data. The results of the other sets are less conclusive.

Looking at the reduced statistic, that is, counts of segments where H_0 could be rejected more than 7 and 9 times for the raw data (64 individual tests) and more than 3 and 5 times for the first differenced data (48 embedding-sets tested), the result for set B1 is confirmed (Fig. 4, bottom row). The dashed lines represent the less demanding statistics (7 and 3 positives requested, respectively), and the continuous line represents the more stringent statistics (9 and 5 positives requested, respectively). Clearly, the non-linearity is relatively labile since few survive this increase in stringency.

Fig. 5 shows the counts of positive embedding-sets over time for all trials of the most promising set, B1. Again, the left column stands for the raw data, the right column for the first differenced data. Each line draws the number of embedding-sets leading to a rejection of H_0. The step height corresponds to 15 counts. There are large variations from trial to trial: compare for example, trial 3 and trial 7, left column. While the cumulative results show an enhanced number of non-linearities during stimulation, the individual trials are less conclusive.
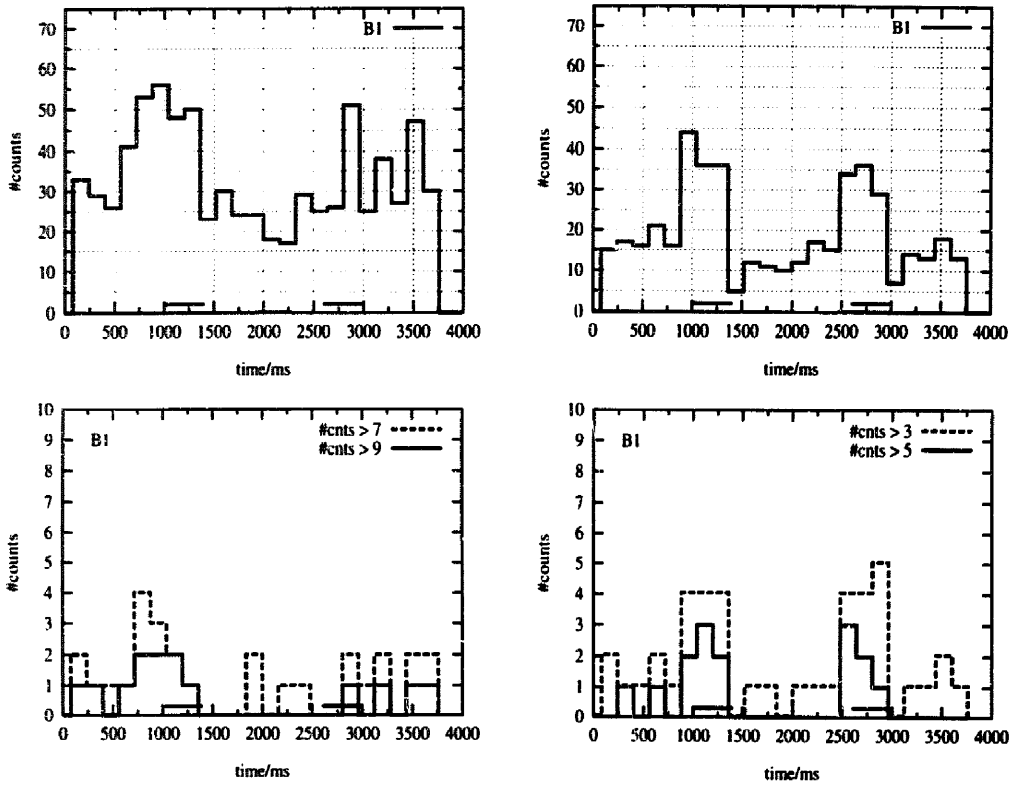
Fig. 4. (top) Tectum FNS cumulative results: left column, original data, right column, first differenced data. Shown here only results from data set B1. Horizontal axis: time. Vertical axis: counts of individual rejections of H_0. (bottom) Tectum FNS reduced statistics. Same conventions as in Fig. 3 for the same data presented in the top row. Vertical axis: number of segments (out of ten) where H_0 could be rejected for more than 7 and 9 different embeddings (original data) and more than 3 and 5 embeddings (first differenced data).

Figs 6 and 7 show in greater detail two trials, trial 5 and trial 7 of set B1. The figures include the original data (field potential and spike activity), the FNS counts for the original data, the first differenced data, and the FNS counts for the first differenced data.

In trial 5 the neural oscillation arising after stimulation is clearly seen, even during the second stimulation where the spike activity is less pronounced. The non-linearity statistics follows roughly the stimulation periods. In contrast, trial 7 shows less clear response to the stimulation, but a comparatively large peak in the FNS statistics at time 2600 ms and 3400 ms. Obviously, there is no robust and simple correspondence between stimulation, cell response and occurrence of non-linearity in these data.

3.2.3. DVS

We applied the DVS method on data segments, where the null of linearly correlated noise has been rejected for at least 5 different embedding-sets. As in the case of the thalamus, we search in the tables of FNS ratios those embedding-sets, where the null has been rejected (maximally one for each embedding dimension). A total of 1035 embedding-sets on 248 data segments has been tested. None of the tests gave support to the hypothesis of low-dimensional dynamics: all DVS plots show the same signature characteristic for high-dimensional dynamics.
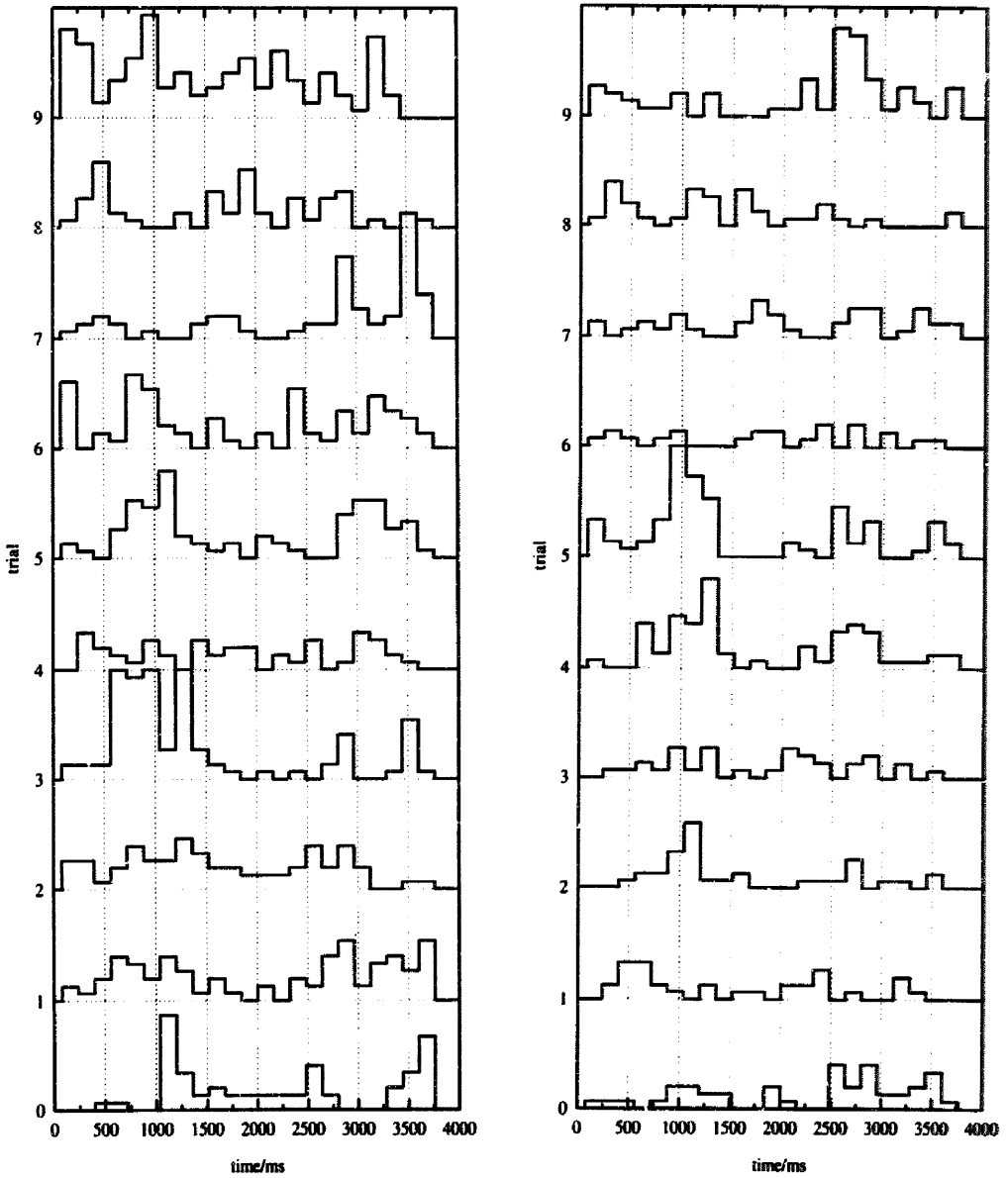
Fig. 5. Tectum FNS trial by trial individual counts for series B1 based on MC statistics. The distance between lines corresponds to 15 counts.

4. Discussion

4.1. The framework

The brain is a distributed dynamical system, comprising billions of interconnected neurons, subject to constant sensori-motor coupling and producing a number of endogenous rhythms and state transitions. It seems unrea-
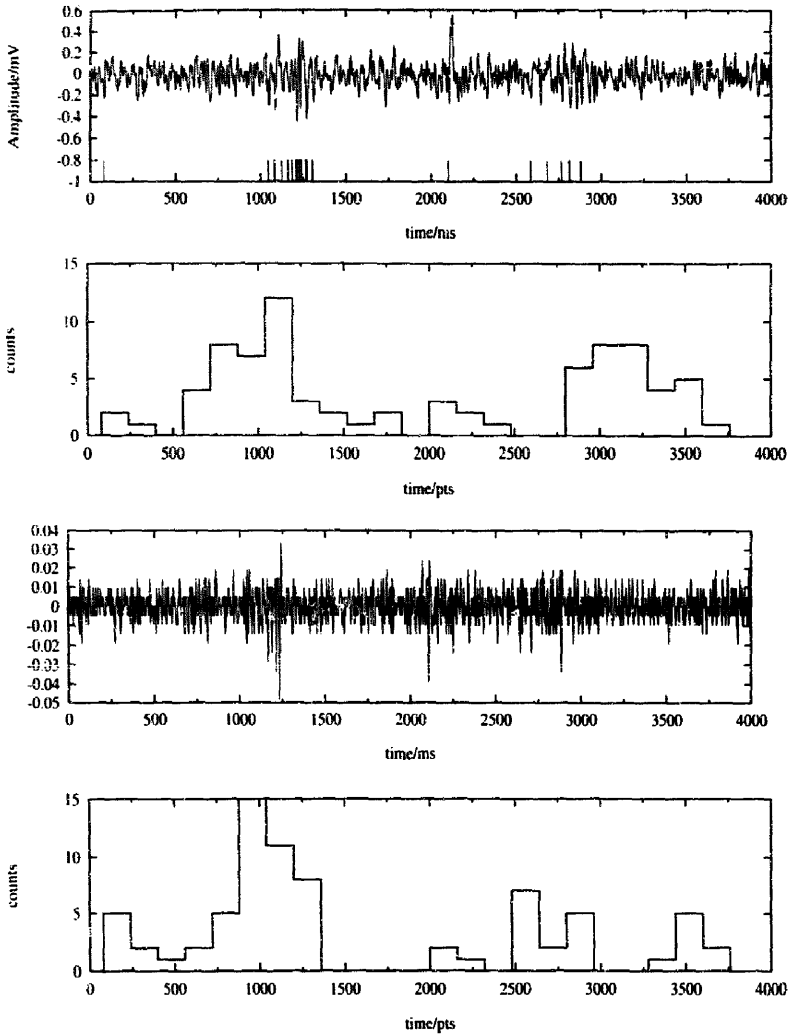
Fig. 6. Tectum recording, series B1, trial 5. From top to bottom: raw data, FNS statistics, first differenced data and FNS statistics for first differenced data. We show a best case of results where both statistics show enhanced significance for non-linearity during the cells response.

sonable to seek for a global, low-dimensional deterministic dynamics in such a situation. We have to clarify here what we mean by *low*-dimensional. If the brain dynamics is confined for some brief time to an attractor of, say, dimension 100, we certainly may call this event low-dimensional, compared to the enormous number of degrees of freedom the complete system possesses. But for the experimentalist this "low" dimensional behavior is not distinguishable from a infinite dimensional one.

The question is not whether the brain is a non-linear system or not; we assume that it definitively is. The question we attempt to answer is whether we can *see* this non-linearity in our recordings and under what conditions, and, further, if we are able to characterize it. In other words, we ask whether the application of non-linear methods is in principle able to provide us with characterizations which are not available through
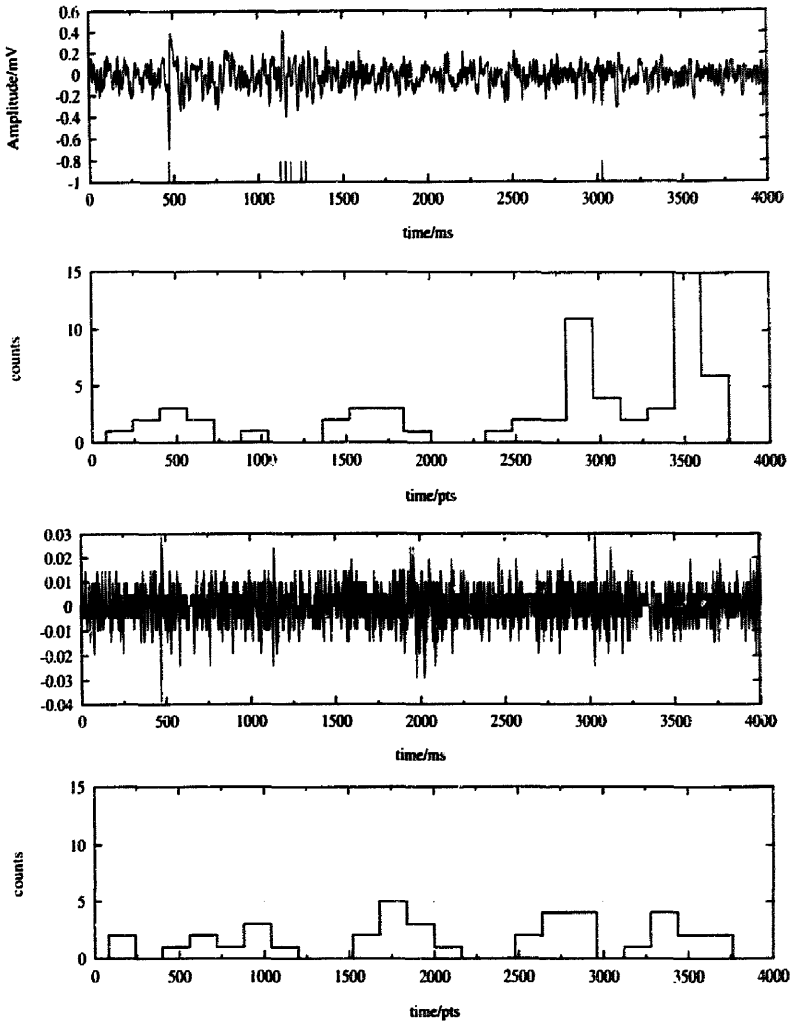
Fig. 7. Tectum recording, series B1, trial 7, showing a poor case of results. Conventions as in Fig. 6.

much simpler and less errorprone linear statistics. Accordingly, our approach here seeks to "reduce" in some way the circumference of the recording site in taking only *local* recordings in the hope to separate a subsystem and to study its dynamics. In addition, we selected the type of signals to be used so as to maximize the chances that they correspond to a cognitive task. Thus we selected visual discrimination responses in visual pathways linked to synchronous neuronal populations. We have a precise knowledge about the onset and duration of the stimulations, as well as the immediate cell responses.

We are now in the position to say that our strategy has been fruitful in that we could submit our time series to detailed statistical scrutiny and come up with reliable answers. As a whole, it can be said that there is evidence for non-linearities in the LFP recordings we have examined. When they appear, they do seem often correlated, as expected, to the onset of the neuronal synchrony that accompanies a visual response. It must also be said that electrical traces obtained under similar conditions did not always give consistent results, and that the traces

of non-linearities were relatively fragile when tested against stringent criteria.

We have observed not one case which displays low-dimensionality. We are led to the conclusion that methods designed to characterize low-dimensional chaotic systems, like dimension analysis or estimation of Lyapunov exponents, are not suited for the analysis of these data. These methods, however, may provide help in the design of empirical tools to uncover further properties of these data.

4.2. Why not more non-linearities?

Why didn't we find, as one might expect under these ideal conditions, robust traces of non-linearity? A first step to answer this question is to notice the discrepancy between the time scales where we expect some coherent dynamics to underly a cognitive task, and the time scales we are forced to use due to the strictures of the analysis procedures. Microscopic cognitive events such as visual discrimination happen on a time scale of around 100–300 ms [48]. The frequencies in the signal are in the range of 5–80 Hz, with the maximal power usually attained around 10 Hz and sometimes around 30 Hz. Reasonable time segments for non-linear analysis methods require at least, say, 5 cycles of the lowest frequency, or 500–1000 ms. In other words, we do not expect a stable dynamics over the time necessary to collect sufficient information about the supposed attractor, and this may well be the cause of its elusive character in our results for the tools at our disposal.

It could be argued that cognitive relevant activity is mainly reflected in higher frequency bands, and declare the lower frequencies to be extra weight in this regard. We could thus try to pursue a search for non-linear structure in higher frequency bands only, and thus use shorter time series. But how can one reliably eliminate the low-frequency "trends" without introducing more artifacts than we wish to discard? One might be tempted to propose adaptive filtering techniques. This, however, is a difficult job, all the more so that there is little detailed knowledge about "first principles" on inner mechanism producing the observed frequencies.

Finally, it is very likely that despite of the control of behavior and the localized nature of the recording sites, the neural dynamics is still too high-dimensional to be seen by a few electrodes.

4.3. Possibilities for future improvements

We have exploited to the best of available techniques a reliable search for non-linear components in single electrode recordings. It has to be emphasized that there are no comparable studies in the literature to date. We have found several possibilities to erroneously classify artifactual non-linearities. While we didn't work on noise reduction methods, we still think our findings relevant: our question was a first order one and since we didn't find any really robust and stable non-linearities, more sophisticated state-space reconstruction techniques should be looked at with skepticism.

What could we do to improve our methods? It is obvious that the brain is an inhomogeneous, spatially extended system, and that we get only scarce information from it when looking only at two electrodes. The obvious alternative is to record from multiple sites, which is known to be a technically laborious neurobiological problem. Further, multivariate non-linear data analysis is far from having reached the state of a standard tool. Today, we simply do not know how to fully exploit multivariate data in a dynamical systems framework. Such basic questions as to whether (1) the recordings on the different electrodes are truly recording new (independent) information, and (2) if the electrodes are recording from the same dynamical sub-system, do not have a satisfying answer.

State-space reconstruction from spatio-temporal signals is still in infancy: to our knowledge there are no results beyond either time-lag reconstruction from scalar time series, or taking multiple sources as basis for the reconstruction space. We have applied this latter approach to obtain global EEG indicators of pathological

conditions with some success, see [28]. But what is needed is a combination of both reconstruction techniques, in order to exploit the dynamical history at each point plus the combined activity together, as in the reconstruction into a *matrix state-space* proposed by [1].

Besides the question of *more* data, we emphasize the need for *better* data. That is, we need extremely clean data, much cleaner than what is obtained today by the competent neurophysiologist who does not confront these issues. Any known noise sources must be minimized before the recording is done. For instance, power line noise might be tolerable within the linear framework, since it is easy to identify spurious frequencies which can be filtered with classical tools. In our Thalamus data we were forced to compromise by filtering a peak of 150 Hz noise which fortunately is largely distant from the range of physiological significance. But in general, in non-linear data analysis there are no a priori unimportant frequencies and no intrusions should be tolerated. All known filter techniques do more harm than good in this respect. This highlights the need for a kind of neurophysiology that conducts experiments with the constraints of such data analysis in mind.

In the long run, obviously, precise techniques for characterizing high-dimensional chaos are going to be necessary for significant advances in this area. Until then, we probably have to content ourselves with less ambitious methods, which are less demanding in data quality and quantity. We have in mind for example spectral estimations of higher order, which provide the first steps beyond linear analysis and therefore more insight in the time series' properties, but are still manageable problems with analytical results.

Acknowledgements

References

[1] H.D.I. Abarbanel, R. Brown, J.J. Sidorowich and L.S. Tsimring, The analysis of observed chaotic data in physical systems, Reviews of Modern Physics 65 (4) (1993) 1331–1392.

[2] H.D.I. Abarbanel and M.B. Kennel, Local false nearest neighbors and dynamical dimensions from observed chaotic data, Phys. Rev. E 47 (5) (1993) 3057–68.

[3] E. Başar, ed., Chaos in Brain Function (Springer Verlag, Berlin, Heidelberg, New-York, Tokyo, 1990).

[4] D. Broomhead and G.P. King, Extracting qualitative dynamics from experimental data, Physica D 20 (1986) 217–236.

[5] M. Casdagli, Chaos and deterministic versus stochastic non-linear modelling, J.R. Stat. Soc. B 54 (2) (1991) 303–328.

[6] M. Casdagli, S. Eubank, J.D. Farmer and J. Gibson, State space reconstruction in the presence of noise, Physica D 51 (1991) 52–98.

[7] M.C. Casdagli and A.S. Weigend, Exploring the continuum between deterministic and stochastic modeling, in: A. Weigend and N. Gershenfeld, eds., Time Series Prediction: Forecasting the Future and Understanding the Past, volume XV of SFI Studies in the Sciences of Complexity (Addison-Wesley, New York, 1993) pp. 347–366.

[8] B. Efron and R. Tibshirani, Bootstraps methods for standard errors, confidence intervals, and other measures of statistical accuracy, Statistical Sciences 1 (1) (1986) 57–77.

[9] J.F. Gibson, J.D. Farmer, M. Casdagli and S. Eubank, An analytic approach to practical state space reconstruction, Physica D 57 (1992) 1–30.

[10] P. Grassberger and I. Procaccia, Measuring the strangeness of strange attractors, Physica D 9 (1983) 189–208.

[11] P. Grassberger, T. Schreiber and C. Schaffrath, Nonlinear time sequence analysis, Int. J. of Bif. and Chaos 1 (3) (1991) 521–547.

[12] M. Hinich, Testing for gaussianity and linearity of a stationary time series, J. Time Ser. Anal. 3 (1982) 169–176.

[13] M.B. Kennel and H.D.I. Abarbanel, False neighbors and false strands: A reliable minimum embedding dimension algorithm, Inst. for Nonlinear Science preprint, UCSD (1994).

[14] M.B. Kennel and S. Isabelle, A method to distinguish possible chaos from colored noise and determine embedding parameters, Phys. Rev. A 46 (1992) 3111–3148.

[15] Y. Kuramoto and I. Nishikawa, Statistical macrodynamics of large dynamical systems. Case of a plane transition in oscillator communities, J. Stat. Phys. 49 (1987) 569–605.

[16] R. Luukkonen, P. Saikkonen and T. Teräsvirta, Testing linearity against smooth transition autoregressive models, Biometrika 75 (3) (1988) 491–499.

[17] R. Luukkonen, P. Saikkonen and T. Teräsvirta, Testing linearity in univariate time series models, Scand. J. Stat. 15 (1988) 161–175.

[18] J. Martinerie, A. Albano, A. Mees and P. Rapp, Mutual information, strange attractors, and the optimal estimation of dimension, Phys. Rev. A 45 (10) (1992) 7058–7064.

[19] F. Mitschke, Acausal filters for chaotic signals, Phys. Rev. A 41 (2) (1990) 1169–1171.

[20] S. Neuenschwander, A.K. Engel, P. König, W. Singer and F.J. Varela, Synchronization of neuronal responses in the optic tectum of awake pigeons, Visual Neuroscience, in press (1995).

[21] S. Neuenschwander and W. Singer, Synchronized oscillations across laminae in the cat lateral geniculate nucleus are stimulus-dependent, Society for Neuroscience Abstracts 20 (1994).

[22] S. Neuenschwander and F.J. Varela, Visually triggered neuronal oscillations in the pigeon: an autocorrelation study of tectal activity, European Journal of Neurosciences 5 (1993) 870–881.

[23] A. Osborne and A. Provenzale, Finite correlation dimension for stochastic systems with power-law spectra, Physica D 35 (1989) 357–381.

[24] A.R. Osborne, A.D. Kirwan Jr., A. Provenzale and L. Bergamasco, A search for chaotic behavior in large and mesoscale motions in the pacific ocean, Physica D 23 (1986) 75–83.

[25] N.H. Packard, J.P. Crutchfield, J.D. Farmer and R.S. Shaw, Geometry from a time series, Phys. Rev. Lett. 45 (9) (1980) 712.

[26] M. Paluš, V. Albrecht and I. Dvořák, Information-theoretic test for nonlinearity in time series, Phys. Lett. A 175 (1993) 203–209.

[27] J.D. Petruccelli and N. Davies, A portmanteau test for self-exciting threshold autoregressive-type nonlinearity in time series, Biometrika 73 (3) (1986) 687–694.

[28] L. Pezard, J. Martinerie, J. Müller, F.J. Varela and B. Renault, Depression as a dynamical disease, Biological Psychiatry, in press (1995).

[29] A. Provenzale, L.A. Smith, R. Vio and G. Murante, Distinguishing between low-dimensional dynamics and randomness in measured time series, Physica D 58 (1992) 31–49.

[30] P. Rapp, A. Albano, T. Schmah and L. Farwell, Filtered noise can mimic low dimensional attractors, Phys. Rev. E 47 (4) (1993) 2289–2297.

[31] P.E. Rapp, A.M. Albano, I.D. Zimmermann and M.A. Jiménez-Monta no, Phase-randomized surrogates can produce spurious identifications of non-random structure, Phys. Lett. A 192 (1994) 27–33.

[32] S.A.R.B. Rombouts, R.W.M. Keunen and C.J. Stam, Investigation of nonlinear structure in multichannel EEG, Phys. Lett. A 202 (1995) 352–358.

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

[34] P. Saikkonen and R. Luukkonen, Lagrange multiplier tests for testing non-linearities in time series models, Scand. J. Stat. 15 (1988) 55–68.

[35] T. Sauer, J.A. Yorke and M. Casdagli, Embeddology, J. Stat. Phys. 65 (3/4) (1991) 579–616.

[36] W. Singer, Synchronization of cortical activity and its putative role in information processing and learning, Annual Review of Physiology 55 (1993) 349–74.

[37] A.C.K. Soong and C.I.J.M. Stuart, Evidence of chaotic dynamics underlying the human alpha-rythm electroencephalogram, Biol. Cybern. 62 (1989) 55–62.

[38] S.H. Strogatz and R.E. Mirollo, Phase-locking and critical phenomena in lattices of coupled nonlinear oscillators with random intrinsic frequencies, Physica D 31 (1988) 143–168.

[39] F. Takens, Detecting strange attractors in turbulence, in: Lecture Notes in Mathematics, D. Rand and L. Young, eds. (Springer, Berlin, 1981).

[40] J. Theiler, Some comments on the correlation dimension of $1/f^\alpha$ noise, Phys. Lett. A 155 (8,9) (1991) 480–493.

[41] J. Theiler, On the evidence for low-dimensional chaos in an epileptic electroencephalogram, Phys. Lett. A 196 (1995) 335–341.

[42] J. Theiler, B. Galdrikian, A. Longtin, S. Eubank and J.D. Farmer, Testing for nonlinearity in time series: the method of surrogate data, Physica D 58 (1992) 77–94.

[43] J. Theiler, B. Galdrikian, A. Longtin, S. Eubank and J.D. Farmer, Using surrogate data to detect nonlinearity in time series, in: Nonlinear Modeling and Forecasting, SFI Studies in the Sciences of Complexity, Vol. XII, M. Casdagli and S. Eubank, eds. (Addison-Wesley, New York, 1992) pp. 163–188.

[44] J. Theiler, P.S. Linsay and D.M. Rubin, Detecting nonlinearities in data with long coherence times, in: Time Series Prediction: Forecasting the Future and Understanding the Past, SFI Studies in the Sciences of Complexity, Vol. XV, A.S. Weigend and N.A. Gershenfeld, eds. (Addison-Wesley, New York, 1993) pp. 429–455.

[45] J. Theiler and P.E. Rapp, Re-examination of the evidence for low-dimensional, nonlinear structure in the human electroencephalogram, preprint (1995).

[46] H. Tong, Non-linear Time Series – A Dynamical System Approach, Oxford Statistical Science Series, Vol. 6 (Oxford University Press, Oxford, 1990).

[47] R.S. Tsay, Nonlinearity tests for time series, Biometrika 72 (2) (1986) 461–466.

[48] F.J. Varela, A. Toro, E.R. John and E.L. Schwartz, Perceptual framing and cortical alpha rythm, Neuropsychologia 19 (1981) 675–686.

[49] H. Whitney, Differentiable manifolds, Ann. Math. 37 (1936) 645–680.

[50] G.U. Yule, On a method of investigating periodicities in disturbed series with special reference to Wolfer's sunspot numbers, Philos. Trans. R. Soc. A 226 (1927) 267–98.

[51] H.P. Zeigler and H.J. Bischof, eds., Vision, Brain and Behavior in Birds (MIT Press, Cambridge, 1993).

[52] S. Zeki, A Vision of the Brain (Blackwell, London, 1994).