

# Characterising the Complexity of Neuronal Interactions

**K.J. Friston, G. Tononi, O. Sporns, and G.M. Edelman**

*Neurosciences Institute, La Jolla, California 92037 (K.J.F., G.T., O.S., G.M.E.); Wellcome Department of Cognitive Neurology, Institute of Neurology, London WC1N 3BG, United Kingdom (K.J.F.)*

---

**Abstract:** This work addresses the complexity of neuronal interactions, the nature of this complexity and how it can be characterised in real neurophysiological processes. A measure of complexity has been introduced recently (Tononi et al. [1994]: Proc Natl Acad Sci USA 91:5033–5037) that is sensitive to the joint constraints imposed by two principles of brain organisation: functional segregation and functional integration. Functional segregation implies that the dynamics of a cortical area should reflect the multidimensional attributes for which that area is specialised (in other words, regional dynamics should show a relatively high entropy). Conversely, functional integration implies a distributed and divergent influence of every cortical area on the remaining areas (i.e., the overall dynamics should show a low entropy). Our measure is based on the profile of entropies of different sized regions of the brain. Complexity is high when smaller regions have (on average) a relatively high entropy with respect to the entropy of the whole system. This measure is equivalent to the (average) mutual information between all small regions and the rest of the system in question.

We have applied this measure to nonlinear simulations and to neurophysiological data obtained with fMRI during photic stimulation. Because patterns of activity in the brain are intermediate between a state of incoherence, with regionally specific dynamics and a state of global coherence, we predicted that simulated nonlinear processes with similar characteristics would have a high complexity. In the language of nonlinear dynamics we hypothesised that the greatest complexity would be found somewhere between high-dimensional, chaotic behaviour and low-dimensional, orderly behaviour. Equivalently, using the metaphor of loosely coupled oscillators, we predicted that complexity would be highest in the domain between asynchronous oscillations and global synchrony. This hypothesis was confirmed using nonlinear neuronal simulations. In addition, we demonstrate that the complexity of neurophysiological data is easily measured and can show a significant complexity when compared to suitable control processes. © 1996 Wiley-Liss, Inc.

**Key words:** complexity, reentry, nonlinear dynamics, information theory, neuronal, functional MRI, visual, functional specialisation, functional integration

---

## INTRODUCTION

The notion that biological systems are found at the interface between chaos and order is a recurrent theme in the sciences of complexity and has some validity when applied to neuronal dynamics [e.g., Fuchs et al., 1992; Gallez and Babloyantz, 1991]. A balance between chaos and order is implicit in the

brain's tendency to diversity, wherein cortical areas preserve their unique and regionally specific dynam-

---

Received for publication January 26, 1995; revision accepted September 26, 1995.

Address reprint requests to Karl J. Friston, Wellcome Department of Cognitive Neurology, c/o MRC Cyclotron Unit, Hammersmith Hospital, 150 DuCane Rd., London W12 ONN, UK.

ics, and in an opposing tendency to integrate regional dynamics into globally coherent patterns of activity. This dialectic is reflected in two aspects of cortical organisation: functional specialisation [e.g., Phillips et al., 1984] and functional integration [e.g., Tononi et al., 1992]. Functional specialisation or segregation is a principle of brain organization that has been best formulated in visual neuroscience [e.g., Zeki, 1990] and identifies cortical areas with a particular function. Specialisation is inferred on the basis of neuroanatomy, connectivity, lesion-deficit studies and, most importantly from the current perspective, selective electrophysiological responses to specific attributes in the visual field. These selective responses distinguish one area from another and suggest that each area expresses its own distinct and separable dynamics. On the other hand, the integration of various perceptual attributes requires functionally segregated areas to influence each other or interact in some way. These reentrant [Edelman, 1993] interactions are mediated by functional or effective connectivity [e.g., Gerstein and Perkel, 1969; Friston et al., 1993] and subserve functional integration. We have proposed a measure of complexity [Tononi et al., 1994] that captures the interplay between functional specialisation and integration. This measure is based on statistical entropies, estimated at different scales or sizes of subsets which comprise a system. In this article we hope to show that this measure can be meaningfully applied to simulated nonlinear systems and to real neurophysiological data.

There are many measures that can be used to assess different aspects of complicated behaviour. Some of these measures are based on the length of a minimal description of the system that is capable of reproducing a set of observations (e.g., Kolmogorov-Chaitin complexity). Measures of this sort reflect some aspect of algorithmic or computational complexity. These measures have a limited role in characterising the nonlinear behaviour of biological systems like the brain. Other complexity measures are derived from information theory (e.g., covariance complexity [Morgner, 1985]) and nonlinear dynamics (e.g., Lyapunov exponents, the correlation dimension or dimensional complexity and metric entropy [Grassberger and Procaccia, 1983]). For example, as the human brain moves between different states of arousal (alert and attentive through to slow wave sleep) the EEG shows characteristic changes, from moderately chaotic (a correlation dimension of six or more) to low-dimensional orderly behaviour (a correlation dimension of about four [Gallez and Babloyantz, 1991]). These measures all

reflect certain aspects of unpredictable or chaotic behaviour, either stochastic or deterministic in nature.

None of the measures above are specifically sensitive to the sort of complexity that characterises neuronal interactions in the brain, namely, a balance between rich chaotic behaviour and coherent orderly behaviour. We focus here on a measure of complexity [Tononi et al., 1994] that is sensitive to both regionally specific dynamics, as suggested by local functional specialisation, and global coherent dynamics, required by functional integration. Only when both of these characteristics are jointly expressed is the system regarded as complex. This measure of complexity was developed in the context of linear systems and is based on the patterns of correlations or covariances among different components of the system in question. This is a useful feature, in that it can be applied to biological data in a practical way that eschews many of the difficulties associated with nonlinear characterisations.

The aim of this paper is to relate our (linear) complexity to nonlinear systems that show complex behaviour. In essence we predicted that as a system is brought from coherent low-dimensional behaviour (where the system is completely integrated but shows no locally specific dynamics) to chaotic high-dimensional behaviour (where the dynamics of each component of the system are easily differentiated, but integration is lost), there would be an intermediate domain of high complexity. In other words, along the continuum of chaos or 'strangeness' of a dynamical system, complexity would show an inverted 'U' type of behaviour. The idea that complexity is high at the interface between non-chaotic and chaotic regimes relates in a clear, if heuristic, way to the notion that (complex) biological systems are typically found at 'the edge of chaos' [e.g., Kauffman, 1992].

At present there are no analytical results pertaining to the relationship between our linear measure and nonlinear measures of chaos (or 'strangeness'). We therefore tested the hypothesis that the greatest complexity, of a simulated nonlinear system, would be found in the region between orderly, low-dimensional behaviour and chaotic, high-dimensional behaviour. Equivalently, in the language of loosely coupled (neuronal-like) oscillators, we predicted that complexity would be highest in the domain of phase transitions between asynchronous behaviour and global synchrony. Phase transitions are abrupt changes in the nature of the attractor manifold as some parameter of the system is changed. In what follows we confirm this hypothesis empirically using nonlinear neuronal simulations. In addition, we show that neurophysiological

data evidence a significant degree of complexity when compared to suitable control processes.

The paper is divided in three sections. The first section describes the complexity measure and its various interpretations. The second section addresses the hypothesis that as a system is moved from chaotic to orderly behaviour, complexity increases, reaches a maximum and then decreases again. This hypothesis was tested by varying the amount of connectivity between simulated neuronal groups. The final section demonstrates that the complexity of real, short and noisy neurophysiological time-series (obtained with fMRI during photic stimulation) can be measured and that the data show a significant complexity.

### THEORY

In this section we review our definition of complexity [Tononi et al., 1994] and its significance for characterising interactions within the brain. The basic idea is that a small functionally specialised area in the brain should have dynamics that contain a lot of information, and that this information is unique to the area in question. On the other hand this specialisation should occur in the context of functional integration, so that information in the area should be available to the rest of the brain. This means that the influence of a small specialised region should be seen elsewhere in the brain. Clearly these two requirements are in conflict; if every small brain region had unique dynamics there would be no integration. If the influence of each area were completely distributed to every other area there would be no specialisation. We propose that a truly complex system is one that manages to resolve this dialectic. A scalar measure that reflects a conjoint expression of functional specialisation and integration is reviewed below. In order to develop the argument in a formal way some technical terms are introduced and explained.

In multidimensional systems entropy and mutual information play the same role as variance and covariance do in univariate processes. Entropy can be usefully pictured as a measure of the volume of the probability density function (p.d.f.) describing the system. (For example,  $m$  scans comprising  $n$  voxels can be plotted as  $m$  points in an  $n$ -dimensional space. The scatter of points will fall mostly in this volume.) A high entropy reflects a large volume where the variables are largely uncorrelated. A low entropy corresponds to a deviation from independence due to substantial correlations between the variables. These correlations (or more generally covariances) distort and reduce the volume of the p.d.f., causing it to collapse along its

principal axes. This reduction in volume can be considered as a measure of the integration of a system [Tononi et al., 1994]. If two multidimensional systems have a high mutual information this means they share a lot of variance or are correlated over a number of dimensions. Even more simply, knowing about one tells us a lot about the other. The entropies ( $H$ ) and mutual information (MI) of two systems ( $\zeta$  and  $\varphi$ ) are directly related: the mutual information is the difference between the sum of entropies of the separate systems and the joint entropy of the systems.

$$MI = H(\zeta) + H(\varphi) - H(\zeta \cap \varphi) \quad (1)$$

where the joint entropy is the sum of the entropy of the first system and the conditional entropy of the second, given the first:

$$H(\zeta \cap \varphi) = H(\zeta) + H(\varphi | \zeta) \quad (2)$$

The conditional entropy [ $H(\varphi | \zeta)$ ] is usually thought of as the uncertainty about system  $\varphi$  that remains after  $\zeta$  is known [Jones, 1979]. A subset of a system is simply a set of some of its constituent variables or components. For example, in a time-series of neurophysiological measurements from  $n$  pixels (functional neuroimaging),  $n$  channels (multichannel EEG recording) or  $n$  separable spike trains (multiunit-electrode recordings), a subset of size  $i$  would represent the time-series from  $i$  pixels, channels or spike trains. We denote the  $k$ th subset of size  $i$  by  $\xi_k^i$ . In what follows we make the simplifying assumption of stationariness, over the time intervals considered.

### Complexity

The key idea behind the concept of complexity considered here is that spatially extended complex systems, such as the brain, show a characteristic profile of entropies over subsets of increasing size: If a system is complex its subsets will have relatively higher entropies than would have been predicted on the basis of the entropy of the whole system. The rationale for this is based on a common sense analysis of functional segregation and integration. Consider a functionally specialised cortical area such as V4. Among other things V4 is specialised for colour. Because colour is a multidimensional attribute (colour cannot be specified using just one dimension or number) the dynamics of any region in a retinotopic map of colour will be multidimensional. In a spatially extended map the dimensionality of the dynamics of all the map's constituent units (e.g., neuronal groups) will increase

in proportion to the resolution and size of the map. This means the neuronal dynamics intrinsic to V4 must span many dimensions (however it is measured) and therefore the entropy of neuronal activity within V4 will be reasonably high. Consider now the implication of integrating many functionally specialised areas like V4. The effect of reentrant interactions between local maps [Edelman, 1978, 1993] will be to distribute the influences of V4 throughout the cortex. The uncertainty (conditional entropy) about dynamics in the remaining cortex will therefore be reduced by measuring the state of V4. Let  $\zeta$  represent a subset corresponding to a functionally specialised area (say V4) and let  $\varphi$  denote the rest of the brain. Functional specialisation implies that  $H(\zeta)$  is reasonably large and integration requires that  $H(\varphi|\zeta)$  be small. The entropy of the whole brain ( $\beta$ ) is given by Equation (2):

$$H(\beta) = H(\zeta) + H(\varphi|\zeta) \quad (3)$$

From Equation (3) it can be concluded that, subject to the constraint of a high  $H(\zeta)$ , integration will be best served by a relatively low entropy on considering the brain as a whole (as low as it can be given that  $H(\zeta)$  is high). Equivalently, under this constraint of a low global entropy, functional specialisation suggests that regional entropies (the entropy of spatially contiguous subsets) will be relatively high.

This characteristic profile of entropies over difference spatial scales is measured in an expedient way by the average entropy for all subsets of size  $i$  [ $\langle H(\xi_k^i) \rangle$ ] minus that which would be expected if the  $n$  components of the system were all independent =  $i.H(\xi^n)/n$  (the latter equality follows from the fact the entropies behave additively for independent systems [Jones, 1979]). The sum of these differences is defined as complexity (C):

$$C = \sum_{i=1}^n (\langle H(\xi_k^i) \rangle - i.H(\xi^n)/n) \quad (4)$$

where  $\xi_k^i$  are subsets of size  $i$  and  $\langle \cdot \rangle$  denotes expectation (over  $k$ ). Note that there is only one subset of size  $n$  ( $\xi^n$ ).

Complexity is therefore simply the area between the two curves of the actual entropy of a subset  $\langle H(\xi_k^i) \rangle$  and that predicted on the basis of the global entropy  $i.H(\xi^n)/n$  as functions of region size ( $i$ ). An example of these curves is seen in the upper part of Figure 1 and the shaded area corresponds to complexity. The upper curve represents the average regional entropies  $\langle H(\xi_k^i) \rangle$  as a function of their size ( $k$ ) and the lower

curve (the straight line) is that which would be expected if all the elements of the system were independent  $i.H(\xi^n)/n$ . It is seen that, given just the entropy of the entire system, the average entropy of smaller regions is much larger than would have been predicted on the basis of independence (i.e., the curved line is higher than the straight line for small subsets). Conversely, given the average regional entropy for a small region, the global entropy is smaller than would have been predicted. This is exactly what would be expected under the joint constraints of functional specialisation and functional integration. Under the constraint of integration (a low global entropy) functional specialisation requires a high regional entropy or conversely, under the constraint of specialisation (high regional entropies) functional integration requires a low global entropy. In Figure 1 the  $\langle H(\xi_k^i) \rangle$  were calculated using a Gaussian autocovariance matrix with width 1.6 and by assuming multivariate normal distributions where:

$$\langle H(\xi_k^i) \rangle = \langle \log([2\pi e]^i \cdot |\gamma_k^i|) / 2 \rangle \quad (5)$$

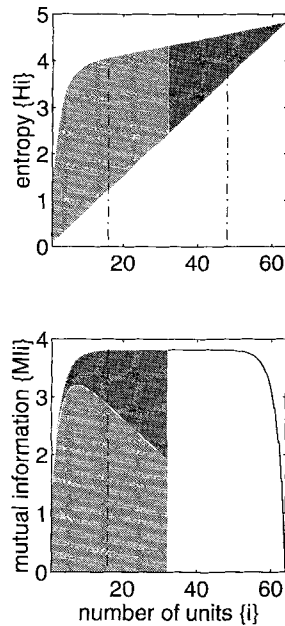
$|\cdot|$  denotes the determinant and  $\gamma_k^i$  is the subset of the covariance matrix corresponding to  $\xi_k^i$ . This is a standard result [see Jones, 1979].

### Complexity and intrinsic mutual information

In this section we consider another interpretation of complexity in terms of mutual information. Whatever the nature of reentrant dynamics, integration suggests that there is necessarily a high mutual information between the dynamics intrinsic to any small cortical region and the rest of the brain. It is easy to show that complexity is the sum of these intrinsic mutual informations. For any subset of size  $i$  there is a complementary subset of size  $n - i$ . The mutual information between these two subsets is given by Equation (1). Let the average mutual information between subsets of size  $i$  ( $\xi_k^i$ ) and their complements ( $\xi_k^{n-i}$ ) be denoted by  $MI^i$ , then:

$$\begin{aligned} MI^i &= \langle H(\xi_k^i) + H(\xi_k^{n-i}) - H(\xi^n) \rangle \\ \sum_{i=1}^{n/2} MI^i &= \sum_{i=1}^{n/2} (\langle H(\xi_k^i) \rangle - i.H(\xi^n)/n) \\ &\quad + (\langle H(\xi_k^{n-i}) \rangle - (n-i).H(\xi^n)/n) \\ &= C \end{aligned} \quad (6)$$

Figure 1 tries to make this relationship clear by showing that the area defining complexity can be split



**Figure 1.**

Graphical illustration of complexity and the relationship between the entropy of regional subsets and the average mutual information between subsets and the rest of the system. **Upper box:** The average entropy of all contiguous subsets of a 64-dimensional process expressed as a function of regional size ( $i$ ). This process was characterised by a stationary Gaussian autocovariance matrix of parameter (width) 1.6. The entropies were calculated according to Equation (5)—upper curve. The straight line is the average entropy of different sized regions that would have been seen if all the 64 components were independent. Complexity is given by the area between these two lines (dark and light shaded areas). **Lower box:** As above but depicting the average mutual information between subsets of a given size ( $i$ ) and the rest of the system. The area under this curve is the same as that defining complexity (dark and light shaded areas). In other words, for all regions of a given size, say 16, the average mutual information between all subsets of size 16 and all subsets of size  $64 - 16 = 48$  is the same as the contributions to complexity from all subsets of size 16 and 48 (broken lines in the upper panel). See main text for a fuller explanation.

into two and added together to give  $MI^i$  (for  $i = 1 \dots n/2$ ). Consider subsets of size 16 (the left line in the top panel of Figure 1). The mutual information between a subset of size 16 and its complement (of size  $64 - 16 = 48$ —the right line in Figure 1) is simply  $H(\xi_k^i) + H(\xi_k^{n-i}) - H(\xi^n)$  for  $i = 16$  and  $n = 64$ . This is the sum of  $H(\xi_k^i) + i.H(\xi^n)$  and  $H(\xi_k^{n-i}) - (n - i).H(\xi^n)$ . If we take the expectation over all  $k$  subsets, then this is the same as the sum of the contribution to complexity by subsets of size 16 and of size 48. These contributions are represented by the light and dark shaded regions in Figure 1 which, when added together (lower panel

in Figure 1), give the average mutual information between subsets of a particular size and their complements. The sum of these mutual informations is again complexity. This equivalence means that a complex system is characterised by a high (average) mutual information between any small (less than  $n/2$ ) subset and the remainder of the system.

A high complexity is consistent with the suggestion that each small cortical area should have a high mutual information with respect to its inputs [see Linsker, 1988; Friston et al., 1992]. More specifically it has been proposed that the entropy of divergent efferent activity from a small cortical region should be high and that the entropy of convergent afferent activity from the rest of the brain should be low. This anti-symmetric arrangement ensures the mutual information between any small region and the rest of the brain is high [Friston et al., 1992] and, from the current perspective, ensures that the interactions are complex.

Like mutual information, complexity is invariant under changes of variables (or units of measurements). For example, the complexity computed on the basis of the covariance matrix is the same as that using the correlation matrix. It should be noted that the subsets implied by 'cortical regions' are anatomically contiguous. In this paper we only consider subsets that are spatially contiguous because of their special relevance to the real brain and functional mapping. It is, of course, possible to compute a whole family of complexities using Equation (4) by selecting the subsets ( $\xi_k^i$ ) in a variety of ways. In Tononi et al. [1994] the sampling was unconstrained and in that paper no reference was made to the spatial contiguity relationships of the elements of the system.

## NONLINEAR SIMULATIONS

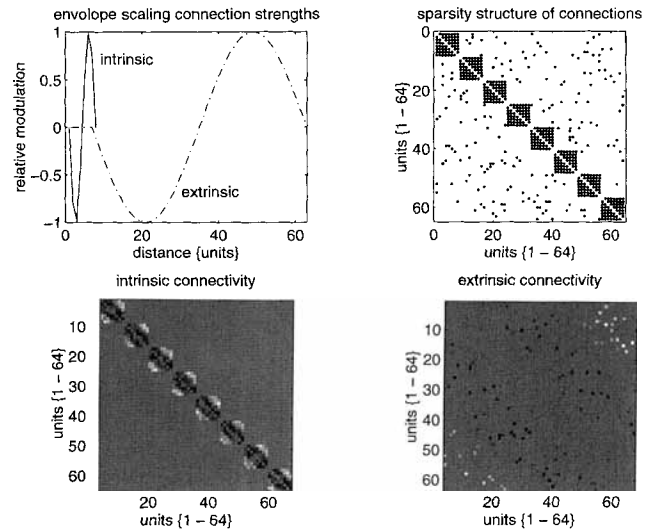
In this section simulations of neural-like processes are analysed using complexity as described above and conventional nonlinear methods. The aim of these analyses was to demonstrate that complexity is highest at some intermediate point, between chaotic behaviour and highly ordered behaviour. Two nonlinear measures were employed to characterise the processes: (1) the correlation dimension ( $D_2$ ), which can be regarded as a measure of the degree of chaos (or strangeness of the attractor [Grassberger and Proaccia, 1983]) and (2) the size of the largest phase-locked cohort [Sporns et al., 1989; Tononi et al., 1992; Lumer and Huberman, 1992]. The former is a standard way of measuring the degree of chaos in a system and can be estimated from the Lyapunov exponents that characterise the system's strange attractor (see Appendix for

further details). The latter measure is meaningful in the context of phasic or oscillatory interactions between neuronal groups and reflects the degree of synchrony within the system.

### Simulations

The simulations were designed to explore a range of behaviours by manipulating just one parameter— $f$ . This parameter changed the relative amounts of two sorts of connections: (1) connections intrinsic to eight groups (of eight units) responsible for coherent intra-group dynamics of reasonably high entropy and (2) sparse extrinsic connections between the groups. The latter connections allowed for intergroup interactions that mediated integrated and synchronous behaviour. By varying the relative amount of intrinsic and extrinsic connectivity the system was brought from complete asynchrony (between groups), wherein each group was uniquely identified by its own dynamics, to near global synchrony, where the eight groups behaved, effectively, as one large group. Figure 2 shows an example of these connections. The mathematical details of these simulations and nonlinear measures will be found in the Appendix.

To illustrate the range of behaviours, elicited by changing  $f$ , three simulations are presented. The simulations were over 2,000 iterations (following 512 iterations to allow for transients) at three values of  $f = 0, 0.17$  and  $0.25$ . The resulting processes are shown in Figures 3, 4 and 5. The first simulation was in the absence of any extrinsic or intergroup interactions ( $f = 0$ ) and shows markedly chaotic behaviour. Figure 3 shows the mean activity over all units as a function of time (top left) and the corresponding phase portrait (top right). The phase portrait depicts the activity of the three largest modes defined by singular value decomposition [Golub and Van Loan, 1991]. The singular value spectrum is seen on the bottom left. This spectrum indicates the number of non-trivial modes (distributed patterns of activity [Friston et al., 1993]) that characterise the system. There were about 16, suggesting each group was contributing about two (topological) dimensions. The profile of  $MI^i$  is shown on the lower right and should be compared with the equivalent profiles from the next two simulations (the area under this profile is complexity). The intrinsic mutual informations were calculated using Equation (5). Figure 4 shows a simulation with substantial, but not an overwhelming amount of extrinsic connectivity ( $f = 0.17$ ). In this case the process is less chaotic and has fewer singular values. The complexity is, however, much greater than in the previous simulation. Al-

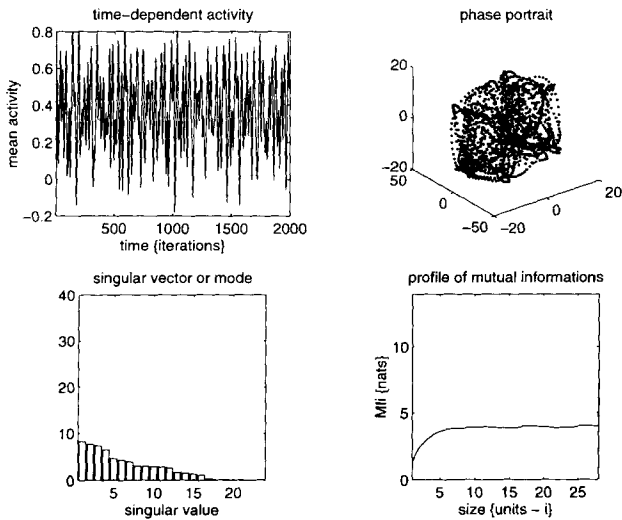


**Figure 2.**

The connection matrices: Each simulation used a single connection matrix that was obtained by combining intrinsic and extrinsic connectivity matrices in varying amounts. **Top left:** Functions describing the modulation of connection strengths (selected at random between 0 and 1) for the intrinsic (solid line) and extrinsic (broken line) connections. These functions are sine waves and simulate local inhibitory and long range excitatory interactions (both within and between simulated groups of units). **Top right:** The combined sparsity structure of both matrices (a dot represents a non-zero connection). Only the extrinsic connections are sparse (5% of all possible connections). **Bottom left:** Intrinsic connection matrix showing the eight groups of eight units. The image brightness reflects the value of the elements of the matrix. The image has been scaled to its maximum. **Bottom right:** Equivalent image of the extrinsic connection matrix.

though the phase portrait in the previous simulation (Fig. 3) may look more ‘complicated’ it is not. This is because the dynamics are not integrated (i.e., the activity of one simulated group predicts nothing about another). Note that processes which appear chaotic or random are not necessarily complex. The third simulation is shown in Figure 5 and shows orderly low-dimensional dynamics due to substantial extrinsic connections ( $f = 0.25$ ). Although the phase portrait suggests that this process is quasiperiodic it is in fact still chaotic. The important thing here is that complexity has fallen.

These three simulations provide anecdotal evidence in support of the hypothesis that complexity is highest somewhere between marked chaos and relative order. To examine this issue systematically  $f$  was increased from 0 to 0.4 in steps of 0.01. For each value of  $f$  the following measures were obtained: (1) the correlation dimension, (2) the size of the largest phase-locked

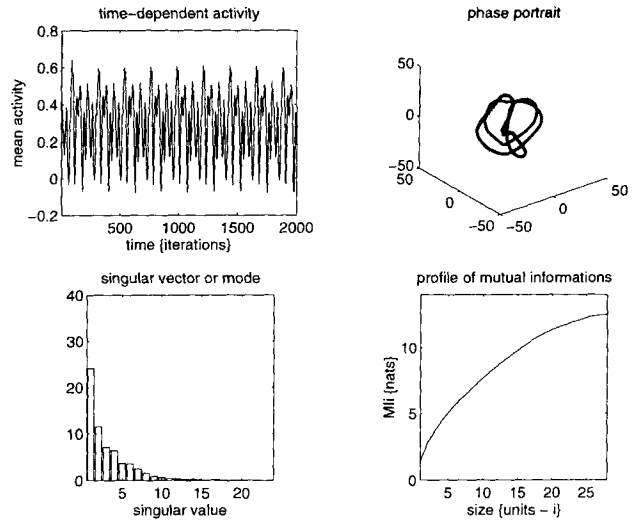


**Figure 3.**

First in a series of three figures characterising simulated processes ranging from chaotic asynchronous behaviour to relatively ordered synchrony. This, the first, was obtained as described in the text using only intrinsic connections. **Top left:** Mean activity ( $\langle x_i \rangle$ ), over all 64 units, expressed as a function of time (iterations). Irregular chaotic behaviour is apparent. **Top right:** A phase portrait, in terms of the activity of the three largest 'modes' identified as the singular vectors with the largest singular values. The space-filling, high dimensional nature of this process is evident. **Bottom left:** The spectrum of singular values following a singular value decomposition. The number of non-trivial singular values reflects the number of 'modes' or distributed patterns of orthogonal activity that characterise the data [Friston et al., 1993]. **Bottom right:** Complexity, shown as the profile of MI<sup>i</sup> over sizes (MI<sup>i</sup> is the average mutual information between all subsets of size  $i$  and the remaining components). Complexity is the area under this curve.

cohort and (3) complexity. The correlation dimension was estimated as described in the Appendix. The size of the largest phase-locked cohort was calculated according to Lumer and Huberman [1992] using the largest Fourier component of each unit. The size of the largest cohort was estimated as the proportion of units in some suitably small frequency range  $[1/(2T)]$ , where  $T$  is the number of iterations. Given that the dynamics of units within a group were coherent, the smallest proportion one would expect to see is about 12.5%.

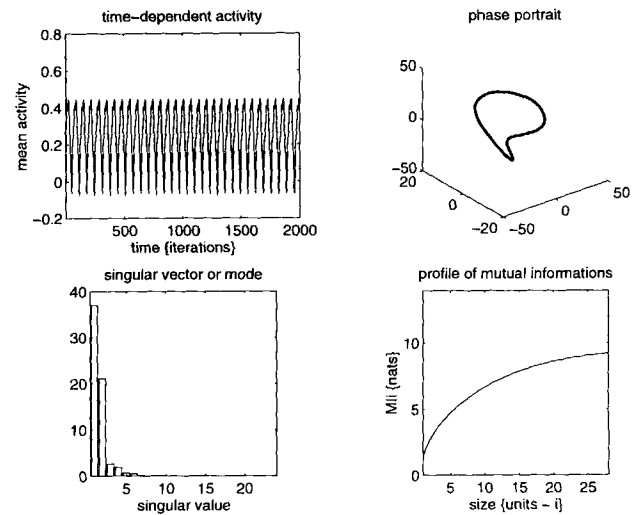
The results of 16 analyses of this sort (averaged) are presented in Figure 6. As the balance of intrinsic and extrinsic connectivity is shifted in favour of extrinsic connections, the correlation dimension falls from very high values ( $D_2 \approx 15$ ) to values just above two (a chaotic process or strange attractor has a correlation dimension of greater than two). Conversely the size of the largest phase-locked cohort increases in an irregu-



**Figure 4.**

The same as for Figure 3 but using a balanced admixture of intrinsic and extrinsic connectivities. It can be seen that the process is more orderly but shows distinct chaotic features. The important thing to note is that complexity for this process is higher than in the previous figure. Complexity is the area under the curve on the lower right.

lar fashion through many phase transitions to reach 70%. Complexity is depicted as the profile of intrinsic MI<sup>i</sup> as a function of  $f$ . The lowest dotted line corresponds to MI<sup>1</sup> and the top line to MI<sup>32</sup>. It is clear that



**Figure 5.**

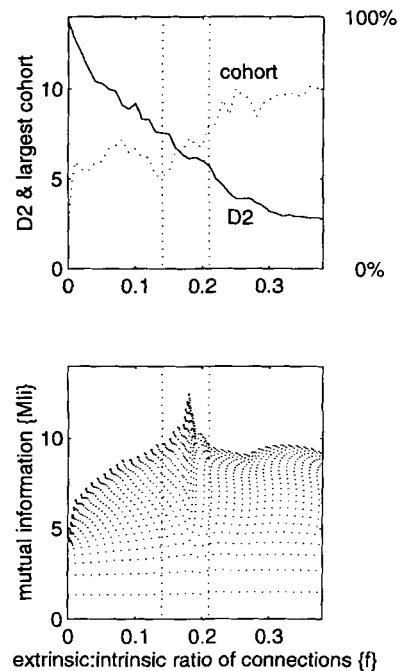
As for the previous two figures but in this instance the extrinsic connectivity dominates. The process is nearly quasiperiodic (the phase portrait shows the attractor manifold to be almost one dimensional) with a smaller number of large singular values. Here the complexity has fallen with respect to the simulation in Figure 4.

complexity is maximal in a limited domain midway between a high and low correlation dimension and during the phase transitions in which the size of the largest cohort is increasing towards maximal levels. We had not anticipated such a 'peaked' profile for the intrinsic mutual informations but this seems to be a robust feature in our simulations. It is interesting to note that the range of correlation dimensions during which complexity is high (vertical dotted lines) agrees very well with the correlation dimensions estimated using EEG in humans [Gallez and Babloyantz, 1991]. However, that the Kaplan-Yorke conjecture used in the estimation procedure (see Appendix) gives only an upper bound on the correlation dimension and this congruence may be only fortuitous.

### THE COMPLEXITY OF fMRI DATA

This section addresses the hypothesis that brain physiology has demonstrable and significant complexity. In general we anticipate that changes in complexity with brain state will constitute most testable hypotheses using complexity. Here, however, the aim is to show complexity is practicably measurable and significantly higher than in a suitable control system. The choice of controls is important because complexity is affected by the nature of the measurements used (e.g., artefactual autocovariances due to poor spatial resolution, noise and low degrees of freedom). In this paper we present two comparisons: (1) between fMRI measurements from within the brain (intracranial) and from the same data but sampled outside the brain (extracranial) and (2) between the intracranial measurements and a simulated fMRI time-series with the same spatiotemporal spectral density (and, implicitly, autocovariance).

The hemodynamic response to sensory stimulation is usually attributed to a physiological uncoupling of regional cerebral perfusion and oxygen metabolism [Fox et al., 1988]. The result is a time-dependent (but uncharacterised) change in the relative amounts of venous oxy- and deoxyhemoglobin. Due to the differential magnetic susceptibility of oxy- and deoxyhemoglobin, a transient change in intra-voxel dephasing is observed. This change subtends the measured fMRI signal [Kwong et al., 1992; Ogawa et al., 1992]. The physiological mechanisms which mediate between neuronal activity and physiology are not, at this time, fully understood; however, the time course of the signal is similar to activity-dependent changes measured with *in vivo* optical imaging, of microcirculatory events, in the visual cortex of monkeys [Frostig et al., 1990].



**Figure 6.**

Results of systematically varying the relative amount of intrinsic and extrinsic connectivities to bring the processes from a chaotic to near-synchronous behaviour. **Top box:** Correlation dimension ( $D_2$ —solid line) and size of the largest phase-locked cohort (broken line) as functions of the parameter  $f$ .  $f$  determines the relative amounts of intrinsic and extrinsic connections and was increased from 0 to 0.4 in steps of 0.01. **Lower box:** Complexity depicted as  $MI^I$  over the range of  $f$ . The top broken line corresponds to  $MI^{I^2}$ . The key thing to note here is that complexity (or  $MI^I$ ) is highest in the domain of phase transitions from chaos to order, or from asynchrony to synchrony. The highest values of complexity were obtained between  $f = 0.1$  and  $0.2$  (dotted vertical lines). The correlation dimension over this range was about 7–4, a range of values found in EEG recordings from the human brain [Gallez and Babloyantz, 1991]. In this domain the largest cohort rose from about 30% to 50% of all the units.

### The data

The data were a time-series of 64 gradient-echo EPI single coronal slices (5 mm thick,  $64 \times 64$  pixels) through the calcarine sulcus and extrastriate areas. Images were obtained every 3 seconds from a single male subject using a 4.0T whole body system, fitted with a small (27 cm diameter) z-gradient coil (TE 25 ms, acquisition time 41 ms). Photic stimulation (at 16 Hz) was provided by goggles fitted with an array of light emitting diodes. The stimulation was off for the first 10 scans (30 seconds), on for the second 10, off for the third, and so on. The data were interpolated from



64 × 64 pixels to 128 × 128 pixels [Keys, 1981]. Each pixel thus represented 1.25 × 1.25 × 5 mm of cerebral tissue. The first four scans were discarded to eschew magnetic saturation effects. Image manipulations and data analysis was performed in Matlab (MathWorks Inc., Sherborn, MD).

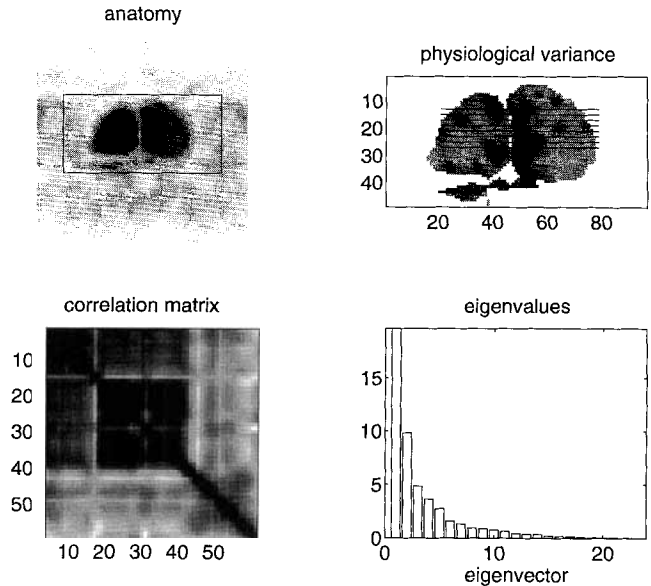
### Analysis

Eight horizontal lines, with a spacing of four pixels (5 mm), were selected (see Fig. 7). Each line segment consisted of 58 pixels. The reason for using these small subsets of the data lies in the limited time-series (64 scans) available. Covariance matrices based on 64 or more pixels would be singular. The control fMRI data were obtained in an identical fashion but using an array of lines outside the brain. The simulated data were created by randomising the phases of the intracranial time-series, following two-dimensional Fourier transform. The results of this manipulation are a stationary time-series with the spectral and autocovariance properties of the original data. Any complexity due to, and only to, poor spatiotemporal resolution will be expressed in these simulated data.

For each of the eight sets of data (real and two control) the complexity was assessed with the profile of  $MI^i$  from one to 28 pixels using Equation (5) and Equation (6). The results of this analysis are shown in Figure 8. It can be seen that the complexity of the intracranial data is significantly higher than both the extracranial and simulated data ( $t = 12.76$  and  $12.65$ ;  $df = 7$ ;  $P < 0.001$ ). It is interesting to note that the extracranial data had higher  $MI^i$  for smaller subsets than the simulated data. Indeed the extracranial data were statistically more complex than the simulated data ( $t = 3.28$ ;  $df = 7$ ;  $P < 0.05$ ). One explanation for this is that during the image reconstruction process neurophysiological variance becomes distributed throughout the image (including extracranial parts). This was confirmed post hoc using singular value decomposition which revealed a substantial extracranial mode time-locked to the periods of photic stimulation.

### CONCLUSION

The dual constraints imposed by functional specialisation and functional integration in the brain suggest that reentrant dynamics should exhibit complexity of a specific nature. A measure of this complexity is now available [Tononi et al., 1994]. One powerful interpretation of this measure is in terms of a high (average) mutual information between local regional dynamics and activity in the rest of the brain. We predicted that

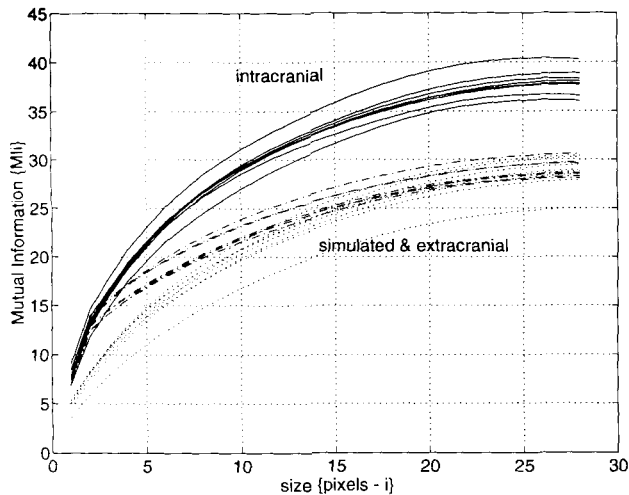


**Figure 7.**

The functional MRI data used in subsequent analyses. **Top left:** MRI scan showing total field of view and the part of the image that corresponds to the map of physiological variance on the top right. **Top right:** An image of the variance in hemodynamic signal over the 64 scans. The image has been scaled to its maximum. The horizontal lines represent the (eight) rows of (58) pixels selected for further analysis. **Bottom left:** The covariance matrix of the time-series from the last row of 58 pixels. Dark means a higher covariance. This section is just above VI according to the atlas of Talairach and Tournoux [1988]. **Bottom right:** Spectrum of eigenvalues of the covariance matrix shown on the bottom left. This spectrum should be compared with the singular value spectra from the simulated processes (Figures 3, 4 and 5).

(1) complexity of this sort would be greatest in simulated nonlinear processes that were similar to the brain (in terms of the balance between high-dimensional chaotic behaviour and low-dimensional orderly behaviour) and (2) complexity would be measurably high in real neurophysiological data. We were able to confirm both predictions.

In this paper complexity reflects a particular form of scaling behaviour, of the entropies measured at different scales of the system in question. Entropy always changes monotonically with the scale or size of the subsystem considered, but in a complex system the rate of change of entropy (with scale) is greater at smaller scales than at larger scales. This allows regional dynamics with relatively rich information (high entropy) to exist in the context of integration at a global scale (low entropy). This scaling behaviour has been framed in terms of functional specialisation and integration of cortical areas; however, we propose that



**Figure 8.**

The complexity in terms of the profile of  $MI^i$  (the mutual information between contiguous subsets of voxels and the remaining voxels) for the eight fMRI time-series depicted in Figure 7—intracranial (solid lines); eight equivalent time-series but extracted from parts of the image outside the brain—extracranial (broken line); and simulated stationary processes with the same spectral and autocovariance properties as the intracranial data (dotted line). As predicted, the intracranial fMRI data show significantly more complexity than the two control data.

this is a general phenomenon that applies at any anatomical scale and to any neurobiological time-series, whether they relate to cortical areas (e.g., cytoarchitectonic regions of interest in fMRI) or not (e.g., multichannel EEG data).

One focus of this paper was to establish a link between information theoretic and nonlinear dynamical characterisations of complexity. This is not as easy as it may seem, given that these two classes of measure pertain to very different aspects of the system or indeed different sorts of system (e.g., stochastic vs. deterministic). Most nonlinear characterisations (e.g., correlation dimension, metric entropy, etc.) can be related to the Lyapunov exponents of the process. The Lyapunov exponents reflect the (average) degree of exponential divergence and convergence of trajectories on the attractor manifold. In this sense they provide a local characterisation of the relationships among the system's variables. Conversely, information theoretic measures (based on the theory of stochastic processes and probability theory) use the global or overall shape of the p.d.f. that ensues when the system is looked at from a stochastic perspective. A deterministic but chaotic process measured with finite and stochastic error would generate a p.d.f. that could be

thought of as a 'blurred' version of the attractor. Therefore, information theoretic measures are comments about the shape of the attractor manifold and regional variations in its density. Of course these two (local and global) characterisations are not independent. For example, the fractal dimension of the attractor manifold will have profound effects on the total amount of space filled by the p.d.f. The first section in this paper provides anecdotal evidence that complexity, based on statistical entropies, can be used to characterise nonlinear systems in a meaningful way. More specifically we demonstrated that there is a systematic and predicted (inverted 'U') relationship between complexity, defined in terms of mutual information, and the degree of chaos, as measured with the correlation dimension.

One interesting link between complexity and the nonlinear dynamics of spatially extended systems is the coexistence of many correlation lengths (distances over which correlations extend) in the vicinity of phase transitions. A heuristic argument could be made that complexity necessarily requires this admixture of correlation lengths: If all the correlation lengths were small this might ensure a high local entropy but would not permit integration. Conversely, if all the correlation lengths were large, integration would be assured (with a low global entropy) but no local specialisation could ensue. Complexity therefore requires a mixture of long and short correlation lengths. We are indebted to Eric Lumer for this observation. In support of this argument our results suggest that, in the region of phase transitions (see Fig. 6), complexity was the highest.

One important difference between nonlinear measures and complexity is a practical one. To reliably measure the Lyapunov exponents (or correlation dimension) one needs considerable amounts of reasonably noise-free, stationary data (some EEG and MEG setups provide data of sufficient quality and quantity to support nonlinear analysis [e.g., Gallez and Babloyantz, 1992]). On the other hand, complexity can be measured with relatively noisy and short time-series, especially if they conform to multivariate Gaussian assumptions. This is important from the point of view of measuring complexity in neurophysiological data from functional imaging and non-stationary electrophysiological time series. The final part of this paper was a simple illustration of how to measure complexity using fMRI. As predicted the complexity of fMRI signals was significantly higher than could be accounted for by spatiotemporal autocovariance or the small number of time points.

The estimation of complexity can be confounded by the nature of the data used. It is easy to show that,

under Gaussian assumptions, convolving the data (with, say, the point spread function of an imaging device) will generally increase the apparent complexity. This follows from the fact that complexity attributable to a convolution is additive and is the complexity of the autocovariances due to the convolution. This complexity is positive for Gaussian convolution kernels. If the point spread function is known it is easy to correct the complexity measure by simply subtracting the complexity due to the resulting autocorrelations. Noise reduces measured complexity. This is simple to see in that noisy data will emulate a system which is more 'independent'. It should also be born in mind that the number of time points can affect the assessment of complexity. This follows from changes in the reliability of the estimated covariances. This phenomenon is encountered when assessing the significance of correlations. The distribution of the correlation coefficient will change with the degrees of freedom associated with the estimate.

These limitations on measuring complexity may seem to challenge its usefulness as a measure. However, as with the correlation dimension, many important neurobiological questions can be formulated in terms of changes in complexity. In these cases the confounding effects of resolution, noise and degrees of freedom can be controlled for by experimental design. In other words, if the resolution of the data, the signal to noise and degrees of freedom of the measured covariances are the same, then the complexity of two systems should be comparable.

A few qualifications deserve mention. First, the results presented in this work are far from exhaustive. We have looked at one very constrained nonlinear simulation and have made no attempt to explore the universe of nonlinear systems either mathematically or empirically. The measure of complexity of the simulations use Gaussian assumptions [Eq. (5)], which are not strictly valid, particularly for low dimensional systems, and therefore the complexity estimates should be regarded as, at best, rough. The empirical determination of entropy, or mutual information, from high-dimensional systems is extremely difficult and at least the methods used here are easily reproduced.

We conclude with a few comments relating to the use of complexity in the analysis of reentrant interactions in the brain. A clear prediction, derived from the relationship between the correlation dimension and complexity in our simulations, is that the complexity of multichannel EEG recording should be maximal in some limited range (probably associated with alert and attentive sensory states). The most powerful prediction here is that the complexity of EEG signals

will show an inverted 'U' behaviour as a function of the correlation dimension. The relationship between the size of the largest cohort and complexity provides an interesting link to the understanding of synchronous neuronal interactions [Eckhorn et al., 1988; Gray and Singer, 1989; Sporns et al., 1989; Tononi et al., 1992]. Simulations exploring phasic interactions in the context of integrating various attributes of visual stimuli have demonstrated a role for distributed phase-locked cohorts of neuronal groups [Sporns et al., 1989; Tononi et al., 1992]. In particular, phase-locked cohorts have been implicated in mediating adaptive responses to conjunctions of disparate visual features. The observation that a high complexity was associated with the formation of large phase-locked cohorts (but not global synchrony) suggests that tasks involving distributed attention may be associated with a higher complexity than equivalent tasks using directed attention. A nice example of this sort of task (performed during functional imaging) can be found in Corbetta et al. [1991], where the 'correct' response depended either on the conjunction of colour, form and motion of target stimuli or on the presence of just one attribute. The degree of integration required by the former task would possibly require a greater degree of phasic reentrant interactions and therefore subtend a greater complexity. With the advent of new functional imaging technologies, these sorts of hypotheses are now testable.

## ACKNOWLEDGMENTS

We are very grateful to Peter Jezzard and Robert Turner at NIH for providing the functional MRI data. We thank our colleagues at the Neurosciences Institute for valuable help and discussion, in particular Eric Lumer and Jonny Wray. This work was carried out as part of the Theoretical Neurobiology Program at the Neurosciences Institute, which is supported by the Neurosciences Research Foundation. The Foundation receives major support for this research from the J.D. and C.T. MacArthur Foundation, the Lucille P. Markey Charitable Trust, and Sandoz Pharmaceutical Corporation. KJF and OS were W.M. Keck Foundation Fellows.

## REFERENCES

- Corbetta M, Miezin FM, Dobmeyer S, Shulman GL, Petersen SE (1991): Selective and divided attention during visual discrimination of shape, colour and speed. *Functional anatomy by positron emission tomography.* *J Neurosci* 11(8):2383-2402.

- Eckhorn R, Bauer R, Jordan W, Brosch M, Kruse W, Munk M, Reitboeck HJ (1988): Coherent oscillations: A mechanism of feature linking in the visual cortex? Multiple electrode and correlation analysis in the cat. *Biol Cybern* 60:121–130.
- Edelman GM (1978): Group selection and phasic reentrant signaling: A theory of higher brain function. In: Edelman GM, Mountcastle VB (eds): *The Mindful Brain*. Cambridge, Mass: MIT Press, pp 55–100.
- Edelman GM (1993): Neural Darwinism: Selection and reentrant signaling in higher brain function. *Neuron* 10:115–125.
- Fox PT, Raichle ME, Mintun MA, Dence C (1988): Nonoxidative glucose consumption during focal physiologic neural activity. *Science* 241:462.
- Friston KJ, Frith C, Passingham RE, Dolan R, Liddle P, Frackowiak RSJ (1992): Entropy and cortical activity: Information theory and PET findings. *Cereb Cortex* 3:259–267.
- Friston KJ, Frith CD, Liddle PF, Frackowiak RSJ (1993): Functional connectivity: The principal component analysis of large (PET) data sets. *J Cereb Blood Flow Metab* 13:5–14.
- Frostig RD, Lieke RR, Ts'o DY, Grindvald A (1990): Cortical functional architecture and local coupling between neuronal activity and the microcirculation revealed by in vivo high resolution optical imaging of intrinsic signals. *Proc Natl Acad Sci USA* 87:6082–6086.
- Fuches A, Kelso JAS, Haken H (1992): Phase transitions in the human brain: Spatial mode dynamics. *Int J Bifurcation and Chaos* 2:917–939.
- Gallez D, Babloyantz A (1991): Predictability of the human EEG: A dynamical approach. *Biol Cybern* 64:381–391.
- Gerstein GL, Perkel DH (1969): Simultaneously recorded trains of action potentials: Analysis and functional interpretation. *Science* 164:828–830.
- Golub GH, Van Loan CF (1991): *Matrix Computations*. 2nd ed. Baltimore and London: Johns Hopkins University Press, pp 241–248.
- Grassberger P, Proaccia D (1983): Measuring the strangeness of strange attractors. *Physica* 9D:189–208.
- Gray CM, Singer W (1989): Stimulus specific neuronal oscillations in orientation columns of cat visual cortex. *Proc Natl Acad Sci USA* 86:1698–1702.
- Jones DS (1979): *Elementary Information Theory*. Oxford: Clarendon Press, p. 152.
- Kaplan J, Yorke J (1979): Chaotic behaviour of multidimensional difference equations. In: Peitgen HO, Walthers HO (eds): *Functional Differential Equations and Approximation of Fixed Points*. Berlin and New York: Springer.
- Kauffman SA (1992): The sciences of complexity and “origins of order.” In: Mittenenthal JE, Baskin AB (eds): *The Principles of Organisation in Organisms*. New York: Addison-Wesley, pp 303–320.
- Keys RG (1981): Cubic convolution interpolation for digital image processing. *IEEE Trans Acoustic, Speech and Signal Processing* 29:1153–1160.
- Kwong KK, Belliveau JW, Chesler DA, Goldberg IE, Weisskoff RM, Poncelet BP, Kennedy DN, Hoppel BE, Cohen MS, Turner R, Cheng HM, Brady TJ, Rosen BR (1992): Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc Natl Acad Sci USA* 89:5675–5679.
- Linsker R (1988): Self organisation in a perceptual network. *Computer (March)*:105–117.
- Lumer E, Huberman BA (1992): Binding hierarchies: A basis for dynamic perceptual grouping. *Neural Computation* 4:341–355.
- Morgera SD (1985): Information theoretic covariance complexity and its relation to pattern recognition. *IEEE Trans on Systems, Man and Cybernetics SMC-15* 5:608–619.
- Ogawa S, Tank DW, Menon R, Ellermann JM, Kim SG, Merkle H, Ugurbil K (1992): Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging. *Proc Natl Acad Sci USA* 89:5951–5955.
- Phillips CG, Zeki S, Barlow HB (1984): Localization of function in the cerebral cortex. Past present and future. *Brain* 107:327–361.
- Sporns O, Gally JA, Reeke GN, Edelman GM (1989): Reentrant signaling among simulated neuronal groups leads to coherence in their oscillatory activity. *Proc Natl Acad Sci USA* 86:7265–7269.
- Talairach P, Tournoux J (1988): *A Stereotactic Coplanar Atlas of the Human Brain*. Stuttgart: Thieme.
- Tononi G, Sporns O, Edelman GM (1992): Reentry and the problem of integrating multiple cortical areas: Simulation of dynamic integration in the visual system. *Cereb Cortex* 2:310–335.
- Tononi G, Sporns O, Edelman GM (1994): A measure for brain complexity: Relating functional segregation and integration in the nervous system. *Proc Natl Acad Sci USA* 91:5033–5037.
- Zeki S (1990): The motion pathways of the visual cortex. In: Blakemore C (ed): *Vision: Coding and Efficiency*. Cambridge: Cambridge University Press, pp 321–345.

## APPENDIX

This appendix presents the details of the nonlinear simulations used in this work and how the correlation dimensions were estimated. Eight groups of eight units each were simulated. The within group or intrinsic connectivities (**I**) were sampled randomly from a uniform distribution in the range (0, 1). They were then modulated by a sine wave envelope, which rendered short range connections negative and longer connections positive (see Fig. 2, upper left). The period of the sinusoidal envelope was equal to the size of the groups (eight units). Matrices of intrinsic connectivity were selected if, and only if, they supported chaotic behaviour. This ensured that each group had its own unique and chaotic dynamics. An identical procedure was used to create the between group or extrinsic connectivity matrix, only in this case the period of the envelope was 64 units. In addition the extrinsic connections were sparse, the sparsity structure being random (5% non-zero elements). The intrinsic (**I**) and extrinsic (**E**) connectivity matrices are shown in the lower part of Figure 2. The intrinsic and extrinsic connectivities were designed to be self-similar so that the any conclusions based on the simulations might ‘scale’. The two matrices were combined according to a control parameter  $f$ :

$$\mathbf{Q} = \sqrt{(1-f)}\mathbf{I} + \sqrt{f}\mathbf{E} \quad (8)$$

This ensured that the Frobenius norm of  $\mathbf{Q}(\sqrt{\sum_{i,j} Q_{ij}^2})$  was constant (=1.5) (both **I** and **E** were scaled such that their Frobenius norms were 1.5). This constraint is fairly arbitrary but ensures that the total amount of

connectivity was constant for each of the simulations. and  
The simulations were obtained by integrating:

$$dx_i/dt = f_i = \phi\left(\sum_j (Q_{ij}x_j + \mu Q_{ij}x_jx_i)\right) - \lambda x_i \quad (9)$$

where  $x_i$  represents the activity of unit  $i$  ( $i = 1 \dots 64$ ).  $\phi$  is the error function (a monotonic increasing nonlinear function). The terms inside the summation represent: (1) a direct influence of unit  $j$  on unit  $i$  and (2) a modulatory, or 'voltage-dependent' effect (scaled by  $\mu = 0.05$ ) which depends on an interaction between afferent input ( $Q_{ij}x_j$ ) and target activity ( $x_i$ ). The final term is simply decay and was sufficiently large to ensure the system was dissipative ( $\lambda = 0.15$ ). Equation (9) was chosen because it is a plausible model of neuronal interactions and at the same time is analytically tractable. The elements of the system's Jacobian are given by:

$$J_{ij} = \partial f_i / \partial x_j = Q_{ij}(1 + \mu x_i) \cdot \phi'\left(\sum_j Q_{ij}x_j(1 + \mu x_i)\right)$$

$$J_{ii} = \partial f_i / \partial x_i = \mu \sum_j Q_{ij}x_j \cdot \phi'\left(\sum_j Q_{ij}x_j(1 + \mu x_i)\right) - \lambda \quad (10)$$

Knowing the Jacobian is important because it allows one to compute the Lyapunov exponents and, from them, the correlation dimension. The Lyapunov exponents were taken to be the expectations of the eigenvalues of the Jacobian [Eq. (10)] evaluated over 256 iterations of the system's trajectory. The Lyapunov exponents reflect the average exponential divergence (and convergence) of nearby trajectories as they course over the local attractor manifold. For the attractor to be strange at least one Lyapunov exponent must be positive (implying divergence in at least one dimension). The correlation dimension of each simulated process was calculated according to the Kaplan-Yorke conjecture using these estimates of the Lyapunov exponents (see Kaplan and Yorke [1979] for further details).